UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THREE PAPERS IN APPLIED MACRO POLICY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ALEXANDER BRENDEN UFIER
Norman, Oklahoma
2015

THREE PAPERS IN APPLIED MACRO POLICY

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF ECONOMICS

BY

_____
Dr. Robin Grier, Chair

_____
Dr. Kevin Grier

_____
Dr. Daniel Hicks

_____
Dr. Gregory Burge

_____
Dr. Scott Linn

Dedication

I dedicate this dissertation to my family and friends, as without their support I would have never finished it.

Acknowledgements

Thanks to committee members Robin Grier, Kevin Grier, Daniel Hicks, Gregory Burge and Scott Linn for their helpful revisions. Special thanks to Ross Hallren for his help with matching. I appreciate the feedback Cynthia Rogers, Souleymane Soumaharo, Li Su, Trey Trosper, Michael Walker and Matt Cobabe provided on various drafts. Thanks to the ETS for having poor website maintenance and never deleting old reports on TOEFL scores which made my third chapter possible, as well as to Michael Keen of the IMF and the research teams at KPMG for their research on the VAT that laid the groundwork for my first and second chapters. Special thanks to Nick Ufier, Stephanie Ufier, Kevin Ufier, and Stephanie Powers for their support through the whole process.

Table of Contents

List of Tables

List of Figures

Abstract

The purpose of any applied macroeconomics study should be to provide actionable information to world governments on policy questions. This dissertation addresses two policies: the first chapter addressed the effects of a VAT on several macroeconomic indicators while revisiting its impact on government balance sheets in the second chapter, while the third chapter considers the effects of English language education.

Value added taxes have become an important source of government funding in past decades, but little empirical work has been done on their macroeconomic impacts. Because the decision to implement a VAT is endogenous and includes variation from all observational units, not just relevant ones, regression methods analyzing the impact of the policy choice will yield biased estimates. To solve this problem, in my first chapter I model the VAT adoption decision for 192 countries using survival analysis. I then match adopters to non-adopters using propensity score matching. I find that VAT adoption is associated with an increase in growth and investment as well as and lower inflation and government spending as a share of GDP. In my second chapter, I use matching techniques to estimate the impact of a VAT on government debts and deficits. The tax is associated with falls in central government debt as well as variance and level of expenditure as a share of GDP.

My third chapter estimates the macroeconomic impact of English skills and considers the utility of English language education. A common language lowers the transaction costs of international trade, and English is increasingly the language of international business. As a result, proficiency with English is often associated with

higher incomes as well as increased employment, trade, and other economic opportunities and is promoted as a policy to improve the wellbeing of people in developed and developing countries alike. However it is not clear whether improved English abilities raise these variables or higher incomes allow one to invest more in better English abilities. Instruction in English is costly and may crowd out other studies or generate other outcomes which negate some of its benefits. I estimate the macroeconomic impact of English skills at the national level, as measured by English proficiency test results from the Test of English as a Foreign Language, TOEFL. I address the endogeneity problem by using the difficulty of learning English given one's native language as an instrumental variable and consistently find a strong effect of English abilities on income and net exports. However, there is no effect of English on FDI or Emigration, suggesting that the impacts of language may come from the changing nature of domestic industries rather than through remittances or foreign investment channels.

The results of these studies should encourage holdout nations to consider adopting a VAT due to its positive effects on macroeconomic growth as well as government balance sheets, and to further explore the advantages of English proficiency. Future research will consider finer points of these policies and optimal implementation of a VAT and English language education.

Keywords: Matching, Survival Analysis, Value Added Tax, English, Instrumental Variables, Linguistic Distance

I. Quasi-Experimental Analysis on the Effects of Adoption of a Value Added Tax (VAT)

1. Introduction

The Value Added Tax (VAT) has displaced many different trade, sales, and manufacturing taxes in the past half century, becoming a major source of government revenue for an increasing number of countries. The tax has all the desirable non-distortionary theoretical properties of a sales tax, its incidence ultimately falling on consumers, but it is much harder to avoid, as the tax is collected at all intermediate steps during the production of a finished retail good rather than at the final sale. As a result, it has proven incredibly popular as it can collect a great deal of revenue (Keen and Lockwood 2010; Toder and Rosenberg 2010). While this tax is theoretically less distorting than other taxes for the each dollar of revenue collected, there is limited analysis of its overall impact on the economy. Demonstrating that the tax also has few negative side effects would make it invaluable for developing countries looking for an effective tax instrument. Showing that a VAT is associated with increases in growth, trade, or investment would be solid evidence in favor of its adoption. Yet empirical studies of the VAT are rare. This paper seeks to fill a hole in the literature, analyzing the overall economic impact of a VAT, which will allow policy makers to make informed decisions of the costs and benefits of this tax.

Understanding the effects of the VAT is vital, as the importance of the VAT is unlikely to decline in the future. The IMF increasingly suggests countries raise more revenue through adopting a VAT or raise VAT rates (Bird 2010). Many countries, facing rising pressures on pension programs due to changing demographics, are funding social programs from general revenues rather than payroll taxes, which usually means

shifting the overall tax composition of government revenues toward VATs (Bird and Smart 2012). A VAT is a requirement for EU membership, and other international organizations are considering their own mandatory VATs, with a unified set of rules and repayment systems at an international rather than national level, such as the Gulf Cooperation Countries (Cnossen 1998; Kapur 2012). As more countries adopt VATs, international payment mechanisms for collecting VATs will become more efficient (Ebrill et al. 2001) and VATs will become more popular in the future.

As a country chooses when it will enact a VAT, simply estimating the effect of adoption on macroeconomic variables, such as trade or growth, through normal regression techniques, will result in biased estimates, as countries that choose to adopt a VAT may be fundamentally different from those who do not. In order to solve this selection problem, I use survival analysis to generate probabilities of VAT adoption to be used as propensity score for 192 countries from 1967-2012. I then employ matching methods to obtain the causal effects of VAT adoption among 112 countries, 90 of which do eventually adopt, from 1986-2007 in the year of adoption and for five years after.[1] This study finds that VAT adoption is associated with increases in investment, and less robustly increases in growth as well as decreases in inflation and government spending. There is no evidence of changes in trade once controlling for selection.[2]

[1] Depending on the variable considered, fewer than 90 treated countries may be compared due to data constraints, but never fewer than 76. See online appendices A-G for more details.

[2] Note that every countries' adoption event will be different, with VATs replacing some taxes in some or being a completely new tax in others. Regardless of the specifics, a VAT will represent a move to shift overall government revenue collection more towards a specific type of consumption tax. This does not imply that it becomes the only, or even dominant form, of taxation revenue. Hence, this paper only can estimate mean effects of adopting a VAT.

This paper is organized as follows. Section II covers background on previous studies on the VAT and other tax policies. Section III describes the adoption equation using survival analysis, both its determinants and results. Section IV introduces the matching process. Section V describes the results of the matching as well as some robustness checks. Section VI concludes.

2. Background

Starting with France in 1954, and reaching its first developing country in 1960 when it was enacted by the Cote D'Ivoire, 166 countries have at some point implemented a VAT as of January 2013, and once passed, nearly all countries retain it. Figure I-1 shows aggregate VAT adoption over time and Figure I-2 shows the adoptions in each year. After slow initial adoption, a number of countries, primarily in Europe and South America, adopt in the mid 1970s. Countries then adopt slowly until the late 1980s, after which adoption rises and then remains high until the early 2000s, with a very large spike following the fall of the Soviet Union. By the mid 2000s adoptions taper off to only a few a year. The general shift towards a consumption type tax that necessarily accompanies the passing of a VAT should have observable effects if different types of taxes have different levels of efficiency or distort economic activities differentially.

Many papers have demonstrated, empirically or theoretically, that consumption taxes like the VAT are superior to income taxes with regard to distortions. Acosta and Yoo (2012) and Arnold (2011) show that countries grew faster when more of their revenue came from consumption taxes. These papers employed panel regression

methods, so selection effects may confound these findings as countries legislate their own tax codes. Krusell et al (1996) create a theoretical model that demonstrates consumption taxes are less distortionary than income taxes, but are less able to achieve redistributive ends. Milesi-Ferretti et al. (1998) finds similar results, with consumption taxes distorting the choice between leisure and consumption less than income taxes for each dollar of revenue raised. It would appear there is some evidence that consumption taxes cause less distortion than income taxes for each dollar raised, but these findings are not specific to VATs and the econometric methods employed may not estimate true causal effects of a given tax policy.

However, there has been less work done to analyze the distortions from VAT adoptions. Keen and Lockwood (2010) study government revenue collected as a result of VATs. Using regression analysis of a system of equations, they found that the higher government revenue was as a share of GDP, the less likely countries were to adopt VATs. VATs generally increased government revenue as a share of GDP, with Sub-Saharan African countries seeing the least benefit. While many authors expect VATs to increase revenues collected, and possibly government spending, due to the relative ease of tax collection, its effects on government consumption and revenues are not clear (Keen 2007; Toder and Rosenberg 2010). Effects on other important macroeconomic variables such as investment, growth, trade and inflation are practically unstudied empirically. While empirical work is scarce, there is a great deal of theoretical work done on VATs that can guide this study.

3. Survival Analysis

3A. Modeling VAT adoption

Understanding the causes and processes of VAT adoption is necessary to address endogeneity due to selection into a VAT. Previously, other authors have modeled the decision to have a VAT as a probit (Keen and Lockwood 2010) or a Markov switching model based on a Cox hazard regression (Cizek et al 2012). Despite their popularity in determining the probability of adopting a policy (Dehejia and Wahba 2002 among others), a panel probit model may not be the best fit for the policy being described in this paper due to VAT's persistence as when a country adopts a VAT, it almost never removes this policy, with only 5 countries ever repealing out of 166 countries adopting (Grandcolas 2005).[3]

A Cox proportional hazard model is employed here instead to model the decision of adopting a VAT. This approach has a number of strengths when applied in this case. First, it fits the decision making accurately: countries tend to adopt a VAT once and never repeal it, so one could model survival time in the pre-VAT state. Second, the hazard model can be used to produce propensity scores for matching. It will predict linear hazards at every period, and two countries with the same predicted linear hazard would be at the same risk of adoption in that period. Thus, among two countries with the same propensity scores, if one adopts and the other does not, one can consider

[3] Vietnam in the 1970s, Grenada in 1986, Ghana in 1995, Malta in 1998, and Belize in 1999 all attempted to pass a VAT and later repealed it. Vietnam, Grenada, and Ghana would later re-adopt a VAT in 1999, 2010, and 1998 respectively, and these re-introduction events that are not repealed are included in this paper. Belize and Malta are included with their original adoption dates. This is in line with what was done by Keen and Lockwood (2010). Omitting these observations does not qualitatively change results.

this assignment to the treatment group as random conditional on observable variables. Third, this selection equation also allows one to estimate the average effects of VAT adoption immediately following the event as well as several years before and after, as adoption is a single event rather than a repeated decision.

Cox proportional hazard models assume all observation units are at risk of failure, in this case adoption of a VAT, at a rate determined by some baseline hazard as well as a vector of time variant and invariant covariates. Only variation from years in which a failure event takes place is used, and once a country fails it contributes no further information to the estimation of the hazard model. More detailed explanation of a Cox proportional hazard model can be found in Cox (1972), with particular attention to the interpretation of covariates and results in Fisher and Lin (1999) and its use in experimental design in Lu (2005).

3B. Determinants of VAT Adoption

The theoretical literature on VATs and taxation comparisons provides a rich pool of covariates that could affect a countries' VAT adoption decision. Among these, international pressure plays possibly the most important role. Keen and Lockwood (2010) propose that countries under IMF lending programs are more likely to adopt a VAT to pay off their debts, both out of necessity for more revenue in general and due to the encouragement they receive from the IMF to adopt a VAT in particular (Bird 2010). This can be seen in some recent IMF reports on the fiscal conditions of many nations; if

a country does not yet have a VAT, the IMF often suggests it should adopt one.[4] I thus include a dummy equal to 1 if a country is under an IMF lending program, either GRA or PRG facility, for any number of months in a given year. Keen and Lockwood (2010) use a similar dummy variable approach. Summary statistics on Table I-1 show that roughly half of the country-years from 1965-2010 are under some sort of IMF lending program.[5]

Another feature of VAT adoption is its general diffusion through proximity. Keen and Lockwood (2010) find evidence that they tend to spread in regional booms and Cizek et al(2012) find significant evidence that adoption may be influenced by a number of spatial factors. Experience with a neighboring country's adoption may better inform a country on the costs of adopting a VAT, or fears of losing manufacturing to a neighbor with a more efficient tax code, could increase the likelihood of a country adopting it. I represent this proximity by creating an index variable that is the equal to share of bordering countries with a VAT, taking a value between 0 and 1. Thus, if a given country borders 4 countries and 2 of them have a VAT, the value of this index is .5 for that country for that year. Table I-1 shows that on average 15% of a countries neighbor's have a VAT, but this measure is heavily skewed right.

Other elements of national economic composition, specifically agriculture and natural resource production, are also likely major determinants of choosing a VAT. Edminston and Fox (2006) found that farmer cooperatives were disproportionately

[4] See IMF Country Report No. 12/54 on Palau, IMF Country Report No. 08/186 on the Marshall islands, IMF Country Report No. 10/191 on Comoros, and the Kingdom of the Netherlands-Aruba: Concluding Statement of the 2010 Article IV Consultation Mission, among others, as well as various newspaper editorials.

[5] IMF lending data is not available before 1984, so this table is only looking at lending data from this point onward.

affected by a VAT, and Cnossen (1998) claims that due to the less formal nature of agriculture, enforcing a complex taxation system is likely to lead to compliance problems. Agricultural production is included as a share of GDP, and it is expected to have a negative effect on the propensity to adopt a VAT. According to Table I-1, on average 23.45% of a countries' GDP comes from agriculture, but the standard deviation is quite high at 17.51%, and so this variable provides a fair amount of variation. Additionally, a country with rich natural resources is less likely to seek tax revenue from other sources (Keen and Lockwood 2010). Keen and Lockwood (2010) use a static measure of subsoil natural resource wealth in 2000 to capture this effect, while instead here I choose to use a series on natural resource rents as a share of GDP from the World Development Indicators. This should accomplish the same objective as the subsoil resource stock in showing the importance of natural resources to a country, but also allow values to change over time, as it is a share of GDP in that year rather than a static stock of natural resources. Resource rents as a share of GDP average 10%, but with a standard deviation of 15.93% the distribution should be quite skewed right.

It is possible geopolitical factors affect the choice to have a VAT. Since authorities need to be able to observe imports in particular, and economic activities in general, in order to tax them, it is possible that large, landlocked countries will have a much harder time collecting a VAT compared to small, island countries due to the number of possible entry locations (Keen 2010). As a result, I have included log of physical size in square kilometers, and whether the country is an island (whether it has no neighboring countries with land borders) or landlocked as potential determinants of VAT adoption. Approximately 20% of country-year pairs are islands and 20% are

landlocked according to Table I-1. On the other hand, as the definition of a small island country is subject to interpretation, I have also included a dummy variable for membership in a political organization, The Alliance of Small Island States, allowing countries to self-select as small islands. While this organization was only founded in 1990, I treat membership as time-invariant, as countries would likely have considered themselves small islands states prior to joining this organization. I also include regional dummies for North America, South America, Europe, Africa, and Asia, with Pacific being the omitted variable.

In addition to time invariant geographic factors, time invariant historical factors may also influence a nation's tax design. Keen (2010) suggests that the British legal system and French legal systems may both, for various reasons, encourage adoption of sales taxes generally and the VAT in particular. Some elements of the VAT in former French colonies bear some resemblance to French tax law, including advanced collection schemes where some sectors are taxed higher than other. It is also possible that the link is not directly from France, but rather just from shared colonial history, with countries with similar colonial pasts cooperating on fiscal policies.[6] British common law countries are more likely to set up independent tax authorities or large tax offices that would make VAT collection more effective with respect to revenue collected, which could make them more willing to set up a VAT in the first place (Keen 2010). To control for this, I include dummy variables for membership or observer status

[6] This reasoning may suggest other colonial dummies may be possible determinants of policy adoption, but no other country colonized nearly as many countries as France or Britain, so adding other colonial dummies will not capture much more variation.

in both the British Commonwealth and the Francophonie.[7] I treat membership as time invariant, so if a country was ever a member of either organization it is treated as a member at all periods for this adoption equation. While membership in these organizations may change over time, often for political reasons, the reason for joining is ultimately a historical one, and thus being a member at any point would indicate some historical influence from the British or French legal systems.[8] According to Table I-1, approximately 30% of country-years are from Francophonie countries and roughly the same amount are from commonwealth countries. Aside from historical origin, the structure of the government may matter as well, specifically whether the country is governed as a federation or not. Federations are less likely to adopt VATs, due to a decentralized tax collecting apparatus (Keen and Lockwood 2010; Ebrill et al 2001; Treisman 2002). However, only 8% of countries are federations.

While not strictly time invariant, membership in the Warsaw pact or being a Former Soviet Republic may also affect VAT adoption. There are a large number of VAT adoptions among such countries in the 1990s after the fall of the Soviet Union. It is thus possible such countries would have adopted a VAT much earlier if they had been allowed to do so by the Soviet Union. Thus I include a time invariant dummy for

[7] Using legal origins gives similar results, as they are highly collinear with membership. However, as legal origin is not a declared membership like membership in the Commonwealth or Francophonie and subject to interpretation.

[8] While membership in the British Commonwealth usually relies on colonial history, the Francophonie is much less restrictive in its membership, allowing countries to join due to more general historical reasons. Several countries in Eastern Europe are members of the Francophonie, for example, despite never being colonized by France.

membership in the Warsaw pact or status as a former Soviet republic.[9] These combined make up only 14% of total country years.

Several other general demographic variables, employed by Keen and Lockwood (2010) in their VAT selection equation, are also included here, including percentage of population that are children below 14 and percentage of population that is above 65 (dependency ratios), log of population level, population growth, log of GDP, and log of GDP per capita. The overall size of the economy and the average wealth of its citizens may affect the decision to adopt a VAT; a very poor country or very small country may not find it worthwhile to construct a large tax collection apparatus that one needs with a VAT. Similarly, the age structure of a country may affect the tax system, as the need for revenue and the ability to collect it changes with demographics (Bird and Smart 2012).

Additionally, the macroeconomic variables likely to be affected by a VAT, investment, growth, trade, inflation, and government consumption, should also be considered as determinants of adoption. Policy makers arguably will take into account the current state of variables most likely to be affected by a VAT; those seeing a low rate of investment or growth may be more likely to adopt a VAT to replace other, more distortionary taxes in hopes of raising investment or growth. Investment, in the form of gross capital formation as a percent of GDP, and growth, as per capita GDP growth, are both included as determinants of the adoption equation.

[9] Former Yugoslavian republics are treated as Warsaw pact countries despite not being members of the Warsaw pact due to sharing similar political cold war history and experiencing turmoil in the 1990's. Due to data demands, these observations often drop out due to missing information for covariates.

Trade in particular is likely to be a major determinant of VAT adoption as well as significantly influenced by VATs. As exports are zero rated for taxation purposes and imports usually full taxed under a VAT (Ebrill et al 2001) countries must consider both the revenue and general economic changes that could arise from such a tax change given their trade patterns (Keen 2008). Hines and Summers (2009) also find that more open countries are more likely to rely on consumption type taxes, such as the VAT. Finally, Swank (2002) shows evidence that increased trade and international competition will affect government taxation behavior. Trade here is represented by exports plus imports as a share of GDP, commonly referred to as openness in the literature.

As VATs have rebate mechanisms in place, they should not lead to increased prices in the same way that other taxes on manufacturing and production do (Ebrill et al 2001). Thus, if a country believes it faces higher than desired inflation, changing its tax system to prevent price cascading is one way that it could affect the price level in the short run. Government spending too is a likely determinant of adopting more taxes, though the sign is of indeterminate direction (Keen and Lockwood 2010). It is possible that individuals in countries with high government consumption as a share of GDP would have high preferences for government services and thus be more inclined to adopt VATs to fund more spending. Alternatively, they may also consider themselves to be at an ideal level of expenditure, and have no desire to raise tax burdens to increase spending further.

Finally, it is important to note the time dimension of adoption. Adoption is defined as the national rollout of the tax system. There is some lag between the decision

to pursue a VAT and the collection of the taxes, both from implementing it politically and creating a tax collection infrastructure. There may also be pilot programs in certain regions before national rollout of the tax, so the effects of a VAT may appear before the tax is national. Thus, I lag all independent variables by two years to account for this, in line with the 18 month span from VAT being introduced in the legislature to collections starting (Ebrill et al 2001). Therefore, if a VAT is adopted in 1988, the determination to do so will be based on values of right hand side variables from 1986.

3C. Results of the VAT Adoption Equation

Table I-2 presents the coefficients (linear hazards) of the Cox hazard regression for selecting a VAT. Five specifications are considered here. The first specification (1) contains all covariates discussed above, and as a result has the fewest number of observations, spanning from 1986 to 2012 with 112 countries, 90 of them adopting a VAT during the observed period. The second specification (2) is the same as the first, but omits government consumption and inflation. The third (3) specification also omits trade, investment, and per capita GDP growth, as these variables are outcomes we will test to see the effects of VAT's adoption, and conditioning the selection on one of the outcome variables may weaken observed effects when one later performs matching.[10] The fourth specification (4) omits many time-varying covariates. This is done in order to study a much longer window where such data is missing and as result observe many more countries, spanning from 1967 to 2012 with 178 total countries and 150 adoptions. The final specification (5) omits all time-varying covariates and looks at only time

[10] I would like to thank an anonymous referee for this suggestion.

invariant factors of adoption, but gains few other observations over specification (4), with 192 countries and 158 adoptions. All countries are considered at risk of adopting a VAT starting in 1960, the year the first developing country adopts a VAT.[11]

Specifications (1)-(3) show no statistically significant differences in marginal effects of covariates on survival. IMF lending has the predicted positive sign, so being under an IMF program is associated with a greater risk of implementing a VAT. Being a member of the Francophonie also sees countries tend to adopt VATs more quickly. Alternatively, having a high share of natural resource rents as a share of GDP is correlated with slower VAT adoption, and countries that are governed as federations also tend to adopt VATs at slower rates. These effects are as predicted in the previous theoretical literature or match the prior empirical findings. No other included variables have a statistically significant effect on adoption rates, aside from continental-level controls.

Specification (4) cannot exactly be compared to (1)-(3) as it contains different covariates, covers a longer period, and has more countries. Being a member of the Francophonie or a Federation have the same effect here as they did in (1), but being a former Warsaw Pact country is associated with a lower likelihood of VAT adoption. This is sensible given the general failure to adopt VATs among these countries until the 1990s. Countries that are larger (as indicated by log population) or have older populations get VATs more rapidly. Specification (5) finds that, upon omitting all time varying variables, larger countries adopt VATs faster and smaller island countries adopt

[11] There is an adoption event in every year from 1984-2012, so specifications 1-3 do not drop any years due to lack of adoption events as would happen with Cox hazard regressions. Specifications 4 and 5 will have some years drop as some years before 1984 have no adoption events.

them slower, along the same effects of being a former Warsaw Pact, Francophonie member or Federal country as seen in specification (4).

4. Matching

The fitted values from the Cox hazard regressions in section III.C can be used as propensity scores. Each country-year pair at risk of adopting a VAT will produce a fitted value *Xb,* the linear hazard of that observation, which describes its current failure rate. Using a simple difference algorithm like that used in Lu (2005), countries with similar hazard rates will have similar expected times of failure, and so any difference in failure time between two country-year pairs Xb_1 and Xb_2 where $|Xb_1\text{-}Xb_2| \leq \varepsilon$ for very small positive ε can be treated as random. A variety of matching methods can be used to compare treatments, but single nearest neighbor matching with replacement is the easiest to discuss and implement. In this method, a country [1] implementing a VAT is matched to a single other country [2] not implementing a VAT based on the choice of country-year pair that minimizes $|Xb_1\text{-}Xb_2|$. This method is used most often in this paper, but other matching methods, kernel matching and radius matching, are also employed to demonstrate robustness of results. The properties of matching estimators are covered in Rosenbaum and Rubin (1983,1985) as well as with a special focus on kernel estimation in Heckman et al (1998).

Figure I-3 provides a visual example of the matching employed here over a general panel of N countries over T years. As time increases from 0 to T, many countries adopt VAT's and region (1) consists of the treatment group country-years for this matching study. After adopting, countries drop out of estimation for the survival

equation, so the country-years in region (2) after adoption are not contributing to the survival equation and not used for matching. As this paper wants to look at effects over time for 5 years after adoption, it cannot use as a control a country 5 years or less before they adopt a VAT, region (3) on Figure I-3. These country-year pairs still contribute information to the survival equation. Countries that do not have a VAT and will not adopt one for 5 years compose the control region (4) and will be used both as controls in the matching and contribute variation to the survival equation. Thus, country-years from the treatment region (1) are compared to similar country-years in the control region (4) to see the effects of adoption.

In order to verify this method has eliminated differences in observables between treatment and controls that would produce bias, one can examine the balance of observable covariates between the matches. One obviously cannot control for unobserved factors in matching. But if matched controls systematically differ from the treated observations on observables, then it will not be clear if the observed difference in outcomes between treatment and control arises from the treatment itself or the underlying differences between the treatment and control groups. Lu (2005) discusses the importance of showing that there exists an equal distribution of all covariates as a result of this problem. Table I-3 is the result of matching all treatment (adoption) country-years to a single nearest neighbor control, then comparing the difference in means for all covariates used in specification (1) of the selection equation from Table I-2. Note that there will be more observations in this check of covariate balance than actually are matched to study given macroeconomic outcomes, as matching looks at a

much longer time window on the outcomes of interest and thus is more sensitive to missing data.[12]

Of the 29 covariates[13] used to determine adoption, one variable shows a statistically significant difference in means between control and treatment at least at a 10% level. Warsaw Pact country-years are statistically significantly more common in the treatment group than in the matched controls. While one may expect a few variables to be significant out of 29 due to random chance, this single difference should not be dismissed out of hand. One could omit all former Warsaw Pact countries as a later robustness check to see if they are driving matching results.

The average propensity score of adopters and matched controls are not statistically different from one another, with a t statistic of .14. Thus the matched control group and adoption group have, on average, the same probability of adopting the policy accounting for the total effect of all observables. However, the average propensity score of adopters compared to all unmatched observations is significantly higher at a 5% level, showing that the adopters, and the controls similar to them, are statistically significantly different from the overall population.[14] These two features are needed for the matching performed to be valid. Another difference of note, while not actually a dependent variable, is the average year difference between controls and

[12] If one checks covariate matches of only those that will be matched for each of the macroeconomic outcomes, the results are qualitatively the same, but this approach requires many more tables, one for each macroeconomic variable (Growth, investment, etc). studied and if one considers matching through multiple methods the number of tables needed grows exponentially.

[13] The number of covariates is 31 if one includes former soviet republics, all of which drop out due to lack of data availability, and the pacific dummy, omitted for perfect collinearity.

[14] Removing 7 very extreme propensity scores among the untreated as outliers raises the t statistic to approximately 9, while it is only 2.22 now.

treatment. The treatment years are on average 4 years later than control years. This is not a surprising result, as there are many more controls that can be drawn from earlier years due to the fact countries exit after adopting in a survival framework. This would be worrisome if earlier years were systematically different from later years; however, as we have already compared 29 variables and found only one of covariates that systematically differ at least at a 10% level (roughly what one would expect for 29 variables), it is unlikely that this difference in years would bias differences in means that is not be captured by these variables.[15]

5. Results

The selection equation used to generate propensity scores for matching is specification (1) of Table I-2, having the fewest observations but containing the most covariates and the highest pseudo R^2 of all specifications attempted.[16] The macroeconomic outcomes of interest here are investment, growth, trade, inflation and government consumption as shares of GDP.

While matching is performed on a single country-year propensity score, the effects on the outcomes of interest of the countries matched are compared several years before and after the event of adoption. This in many ways resembles the method employed by Abadie and Gardeazabal (2003) in that it matches two countries (or more generally in their method, one country to a synthetic composite country that can be composed of one or more countries) then follows their respective paths over time, both

[15] I cannot dismiss the possibility that years do matter for the outcomes, but it would have to work through a channel not captured by the current observables.
[16] Evaluating through SIC and AIC is indeterminate between which specification, 1 or 2, is superior, but specification 1 allows for analysis of the effects of more variables.

forward and backward, to identify the impact of a policy change. The countries should be identical before the policy change and different after if the policy had a causal effect. In the case of this paper, the decision to adopt occurs at time t=-2 and adoption occurs at time t=0. I will consider looking back as far as time t=-7, 5 years before the decision to adopt occurs, and as far forward as time t=5, 5 years after adoption. Furthermore, I will also consider the change in outcomes of interest from 7 years before adoption to 2 years before adoption, when the decision to adopt a VAT is made, as well as the change from 2 years before adoption to 5 years after adoption, to capture the impact over time of VAT adoption.[17]

As one seeks to observe the average treatment on the treated (ATT), control and treatment countries should not differ in the years before a country decides to adopt a VAT. The control and treatment groups should have the same means of the variable of interest before the VAT is implemented, and the between-group difference in changes between 7 years before adoption to 2 years before should be zero. Thus, while both control and treatment group may see changes over time in the variable of interest, these changes should be the same. However, if the implementation of a VAT has an effect, then there should be a difference between changes in the treated group and the control group over time, 2 years before adoption to 5 years after, that is not zero. Thus while the variable of interest may change for both the control and treatment group, the changes should be different between treatment and control groups to assert a causal effect of the

[17] Because actors may expect a VAT, they may change their behavior before it is actually being collected, so looking at the effect from decision to adopt forward may provide a better picture of the overall change rather than starting 2 years later, when it is implemented. Pilot programs that implement a VAT in certain regions a few years before national level implementation may mean a country sees some smaller effect prior to nationwide adoption.

policy. Also, if the average difference in means between matched controls and treatment is nonzero in a given year post-adoption, this is further evidence of an effect.

Results from matching, compared treated to matched controls, are summarized by their t statistics on Table I-4 for the five variables of interest over the thirteen year periods being tested, seven years before adoption to five years after adoption. Most importantly, the average differences in changes over time before adoption as well as after adoption are listed at the bottom of the table. Results comparing treated to unmatched controls, the whole data set without accounting for selection, are also presented for comparison purposes.

The overall differences in changes over time after adoption are statistically significantly different from zero for investment, being positive at a 10% level, and inflation and government spending, being negative at a 10% level. Growth is statistically significantly higher in the year of adoption, along with three and four years after adoption, and inflation is lower four and five years after adoption. These results would be consistent with VATs raising growth and investment as a share of GDP at the same time they lower inflation and government spending. Combining these generally positive effects with previous author's results that VAT can collect a great of revenue, a VAT seems an attractive policy.

Also important for the validity of the results, the overall difference in changes before adoption between adopters and controls is not statistically different from zero for all five variables. There is also no difference in any of the given variables in any of the years before adoption. If there were a differences, or differences in changes, of any variable, it would call into question the interpretation of the results, as it would imply a

difference in trajectories between controls and treatment prior to undertaking the treatment. As there is no such difference in trajectories before, and a difference in trajectories after, there is evidence that VAT adoption caused changes over time among these variables.

Discussing specific variables, it would thus appear over time after adoption, VATs caused a gain in investment as a share of GDP. Ferretti and Roubini (1998) forecast that investment should rise when taxes are collected more from consumption relative to income, which is borne out here. This finding would seem to disagree with Cai and Harrison (2011) who found that Chinese firms did not change investment behavior in response to temporary changes in the VAT rate. However, as Cai and Harrison studied repeal of VATs for firms, not imposition, and looked at a firm rather than a national level, a number of reasons that could explain the difference in findings.

There is no overall difference in changes in GDP growth either before or after the VAT is passed, but several years after adoption growth is statistically significantly higher in countries that adopt compared to similar countries who did not. This would seem to demonstrate against any large efficiency losses as forecast by Stiglitz and Shahe (2005) and Stiglitz (2008) due to the shifting of production from the observed formal sector to the unobserved, less efficient informal sector. If VAT is causing distortions that drive economic activity underground, which should show up in reduced GDP growth numbers, the distortions such as increased evasion efforts are either too small to detect or are offset by GDP gains. Positive changes in growth would agree with the proposed efficiency gains, as would be forecast by Zagler and Durnecker, (2003);

Arnold, (2011); and Hines and Summers, (2009) from moving toward consumption taxes. However, it is unlikely that this higher growth would continue indefinitely.

There is no statistically significant effect of VAT adoption on trade as a share of GDP as seen in Table I-4. Before adoption, trade as a share of GDP is the same among adopters and matched non-adopters. Compared to all countries, the unmatched controls, countries that would adopt a VAT traded less. After adopting a VAT, while adopters do see trade increase as a share of GDP, so do the matched controls. Thus there does not appear to be a causal effect of VATs on trade. However, compared to unmatched controls, after adopting a VAT, countries no longer trade less than other countries.

These results are not inconsistent with the results of Hines and Desai (2005) who found countries with a VAT traded less. While it is true countries that adopt VATs trade less than countries that do not adopt VATs, when matched against countries with similar characteristics, they do not trade any more or less. Even after adoption, when countries with VATs see their trade numbers rise, they still trade (albeit statistically insignificantly) less than unmatched controls. Thus panel methods here could find a negative correlation of VAT adoption with trade. However, this may be because countries who adopt a VAT trade less to begin with, and they may adopt VATs in order to attempt to raise trade; VATs themselves do not lower trade.

Due to wide ranges of possible values for inflation, and its high variance, matching is particularly useful in determining the effects of VATs on inflation. While before adoption countries with VATs tended to have roughly the same inflation as matched controls, in the years following adoption countries with a VAT see a less

inflation in several years compared to matched controls. Due to large yearly swings in average inflation, this result must be interpreted with caution.

The changes in government consumption as a share of GDP prove to be the most interesting. First, compared to unmatched countries, VAT adopters have significantly smaller government spending as a share of GDP, but not when compared to matched countries. There is no difference in spending in any given year before or after adoption, and no difference in changes before adoption, but following the decision to adopt there is a statistically significant decrease over time in government consumption as a share of GDP among countries with a VAT compared to matched controls. As Keen and Lockwood (2010) found sometimes conflicting results of the effects of VATs on government revenue, this result is not unexpected. If collecting public funds and as a result government consumption has become less distortionary due to a more efficient tax system, one would expect government consumption to grow rather than due to the lower marginal cost of collecting public funds makes spending easier. However, if government changes other taxation or spending behaviors at the same time, such as lowering taxes collected by more distortionary means, that could explain this net fall in spending.

It is possible that the earlier results from VAT adoption on various variables arises from the choice of matching method, so six other methods are tested; single nearest neighbor without replacement, nearest three neighbors with replacement, caliper matching with a .025 and .05 radius, and kernel matching with an Epanechnikov as well as Normal kernel. Results are shown in Table I-5, comparing differences in changes over time before adoption and after adoption. Results for investment as well as

government spending and inflation are reasonably robust. On the other hand, growth is less robust, although one may expect single higher years of growth rather than long term growth. The impact of VAT adoption on trade is not statistically significantly different from zero. Finally, before adoption, there is no systematic difference in changes between control and treatment groups on any variable regardless of matching method, which one must see for the inferences above to be valid. Table I-5 provides strong evidence these changes observed after adoptions are causal effects of the policy and not driven by choice of matching methods.

6. Conclusion

VATs are believed to be able to raise a large amount of money with little overhead or distortion, and so have become more common around the world in recent decades. However, studies actually analyzing the impacts of VATs are rare and their results are often not conclusive. Resolving this lack of conclusive results is important, as signs are that VATs and consumption taxes are likely to grow in importance for government funding in future years (Hines and Summers 2009; Bird 2012). This policy deserves more analysis before rushing headlong into more VAT reliance.

As VATs are adopted voluntarily by nations, it can be hard to obtain causal estimates of the effects of the policy due to endogeneity, as those who choose the tax may fundamentally differ from those who did not. This paper proposes solving the problem through the use a of a two-stage process. In the first stage, country's decisions to adopt a VAT are modeled as a Cox hazard regression, with countries surviving for a number of years, based on a vector of covariates, before adopting a VAT. Countries'

survival times are random conditional on the linear hazards from the survival, so any difference in times of VAT adoption is random between two country-year pairs with sufficiently close fitted values. In the second stage, each country-year pair in which an adoption event occurs is matched to a country-year pair in which it does not occur with the closest linear hazards. Then, I compare the difference in means for investment, growth, trade, inflation and government spending between treated and controls at time of adoption and several years after the adoption event, comparing outcomes before the event as a falsification test. There is evidence that VATs lower inflation and government spending, as well as raise investment and growth. Countries most likely to adopt VATs trade less as a share of GDP than countries as a whole, and those that adopt VATs do see a statistically sigificant gain in trade, but no larger than those of matched controls that do not adopt the treatment.

These findings can be of use to many policymakers considering implementing a VAT. VATs carry with them no major side effects in the form of lowering trade and may raise growth and investment, with no feared adverse effects on inflation or government consumption. As these results are the mean effects of VAT adoption, they leave the door open to investigating more detailed aspects of VAT adoption. This study only includes VAT adoption events in over roughly 2 decades, and many earlier VAT adopters may have had systematically different experiences, and future ones may see different impacts as well. This study does not keep track of VAT rates themselves, how many different VAT categories are used by the country, the minimum firm size to register for the VAT, or the speed at which one receives rebates. All of these are potentially major determinants of the ultimate effect of a VAT which could prove

important to the actual effect of VAT adoption (Edminston and Fox 2006). Finally, many VATs did not arise in a vacuum but rather took their place among some existing taxes and perhaps replaced others. Thus, the true impact of any VAT is likely dependent on a rich set of nation-specific tax and other factors. Rather, this paper provides a jumping off point for empirical studies of finer points of the VAT, while at the same time applying a new method to solve important unanswered questions about the empirics of this popular tax policy.

Table I-1
Summary Statistics, Right Hand Side Variables, 1965-2010

Variable	Obs.	Mean	Std. Dev.	Min	Max
Government Consumption	3692	14.22	10.86	0.64	69.83
Inflation	3357	29.01	488.18	-49.42	26762.02
Investment	3060	22.89	10.53	-17.40	113.58
GDP per Capita Growth	3350	1.91	7.12	-50.29	92.59
Trade	3213	78.72	53.33	0.31	432.30
IMF Lending Dummy	1929	0.49	0.50	0	1
Agriculture	2861	23.45	17.51	0	93.98
GDP, Log	3364	22.19	2.31	16.13	30.09
GDP Per Capita, Log	3360	7.32	1.54	4.00	11.02
Natural Resource Rents	3003	10.00	15.93	0.00	155.22
Bordering Countries with VAT	4823	0.15	0.29	0	1
% of Population 0-14	4425	38.25	8.81	11.51	53.03
% of Population 65+	4425	4.87	2.94	0.43	14.63
Population Growth	4790	2.10	1.71	-11.00	18.59
Population, Log	4791	14.75	2.18	8.83	20.88
Size, Log Square Kilometers	4823	10.93	2.87	0.69	16.61
Alliance of Small Island States	4823	0.23	0.42	0	1
Former Soviet Republic	4823	0.07	0.25	0	1
Former Warsaw Pact	4823	0.07	0.26	0	1
British Commonwealth	4823	0.33	0.47	0	1
Francophonie	4823	0.28	0.45	0	1
Island	4823	0.24	0.43	0	1
Landlocked	4823	0.22	0.42	0	1
Federal	4823	0.08	0.27	0	1
Asia	4823	0.30	0.46	0	1
Europe	4823	0.16	0.37	0	1
Africa	4823	0.31	0.46	0	1
North America	4823	0.11	0.32	0	1
South America	4823	0.04	0.19	0	1
Pacific	4823	0.08	0.27	0	1

Notes: **Observational units are country-year pairs.**
Variables are lagged two periods

Table I-2
Cox Hazard Regressions for VAT Adoption

	(1) 1986-2012	(2) 1986-2012	(3) 1986-2012	(4) 1967-2012	(5) 1967-2012
Government Consumption	-0.022				
	(0.017)				
Inflation	-0.007				
	(0.005)				
Investment	0.011	0.011			
	(0.015)	(0.014)			
GDP per Capita Growth	-0.010	-0.002			
	(0.019)	(0.018)			
Trade	-0.003	-0.002			
	(0.004)	(0.004)			
IMF Lending Dummy	1.253***	1.293***	1.302***		
	(0.360)	(0.347)	(0.342)		
Agriculture	-0.008	-0.005	-0.009		
	(0.014)	(0.013)	(0.012)		
GDP, Log	6.359	6.453	6.566		
	(4.505)	(4.184)	(4.246)		
GDP per Capita, Log	-5.968	-5.953	-6.106		
	(4.491)	(4.172)	(4.236)		
Natural Resource Rents	-0.030**	-0.034**	-0.035**		
	(0.014)	(0.014)	(0.013)		
Bordering Countries with VAT	0.512	0.589	0.608	-0.071	
	(0.450)	(0.441)	(0.428)	(0.324)	
% of Population 0-14	-0.004	0.000	-0.005	0.038	
	(0.042)	(0.040)	(0.039)	(0.024)	
% of Population 65+	-0.043	-0.017	0.002	0.152**	
	(0.112)	(0.108)	(0.106)	(0.073)	
Population Growth	0.004	0.011	0.012	-0.017	
	(0.090)	(0.086)	(0.086)	(0.069)	
Population, Log	-5.930	-5.978	-6.066	0.362***	
	(4.491)	(4.171)	(4.236)	(0.094)	
Size, Log of Square Kilometers	-0.032	-0.071	-0.078	-0.095	0.133***
	(0.106)	(0.103)	(0.103)	(0.071)	(0.046)
Alliance of Small Island States	-0.372	-0.419	-0.393	-0.674	-1.085***
	(0.587)	(0.551)	(0.534)	(0.412)	(0.403)

Former Soviet Republic	-	0.368	0.369	0.245	0.409
	-	(1.281)	(1.214)	(0.407)	(0.356)
Former Warsaw Pact	-0.953	-1.109	-1.290*	-0.958**	-0.679*
	(0.716)	(0.692)	(0.687)	(0.432)	(0.374)
British Commonwealth	0.005	-0.001	-0.012	0.303	0.308
	(0.316)	(0.317)	(0.311)	(0.226)	(0.220)
Francophonie	0.681**	0.683**	0.685**	0.487**	0.404**
	(0.300)	(0.293)	(0.290)	(0.206)	(0.196)
Island	0.685	0.794*	0.780	0.012	0.189
	(0.484)	(0.477)	(0.476)	(0.388)	(0.347)
Landlocked	0.076	0.117	0.143	0.049	-0.092
	(0.321)	(0.303)	(0.298)	(0.214)	(0.204)
Federal	-2.056***	-2.184***	-2.201***	-1.047***	-0.695**
	(0.654)	(0.649)	(0.649)	(0.374)	(0.349)
N	1287	1317	1400	4421	4823
Pseudo-R^2	0.118	0.112	0.117	0.071	0.051
Log Likelihood	-313.114	-328.122	-334.086	-635.170	-697.526
Chi Squared	83.686	82.685	88.906	96.815	74.938
Subjects	112	113	114	178	192
Adoptions	90	95	96	150	158

Standard Errors statistics in parentheses.

*** $p < .1$, ** $p < .05$, *** $p < .01$**

Notes: **Observational units are country-year pairs; Values are coefficients for linear hazards, not hazard ratios. A coefficient of 0 thus implies no effect. To find the hazard ratios, exponentiate the coefficient; All independent variables are lagged 2 periods; All estimates use regional dummies for North America, South America, Europe, Asia, and Africa (Pacific omitted.)**

Table I-3
Balance of Covariates for Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Government Consumption	Unmatched	11.82	15.03	-3.21	1.18	-2.71***
	ATT	11.82	9.84	1.98	1.32	1.5
Inflation	Unmatched	12.20	77.66	-65.46	105.71	-0.62
	ATT	12.20	11.11	1.09	4.07	0.27
Gross Capital Formation	Unmatched	22.21	22.73	-0.52	1.25	-0.41
	ATT	22.21	22.23	-0.02	1.78	-0.01
GDP per Capita Growth	Unmatched	1.63	1.63	0.00	0.73	0.00
	ATT	1.63	1.06	0.57	0.92	0.61
Trade	Unmatched	74.66	84.43	-9.77	5.50	-1.78*
	ATT	74.66	62.99	11.67	8.07	1.45
IMF Lending Dummy	Unmatched	0.72	0.47	0.26	0.06	4.63***
	ATT	0.72	0.69	0.03	0.08	0.4
Agriculture	Unmatched	22.47	22.38	0.09	1.91	0.05
	ATT	22.47	23.78	-1.31	3.04	-0.43
GDP, Log	Unmatched	22.66	22.04	0.63	0.26	2.41***
	ATT	22.66	23.27	-0.60	0.47	-1.29
GDP Per Capita, Log	Unmatched	7.07	7.10	-0.03	0.17	-0.19
	ATT	7.07	7.14	-0.07	0.31	-0.24
Natural Resource Rents	Unmatched	6.76	11.82	-5.06	1.61	-3.14***
	ATT	6.76	7.92	-1.16	1.94	-0.6
Bordering Countries with	Unmatched	0.38	0.21	0.17	0.04	4.60***
	ATT	0.38	0.28	0.10	0.07	1.43
% of Population 0-14	Unmatched	35.99	39.83	-3.84	0.85	-4.52***
	ATT	35.99	36.53	-0.55	1.72	-0.32
% of Population 65+	Unmatched	5.62	4.14	1.48	0.27	5.41***
	ATT	5.62	5.53	0.09	0.64	0.15
Population Growth	Unmatched	1.82	2.17	-0.35	0.17	-2.12**
	ATT	1.82	2.01	-0.19	0.23	-0.8
Population, Log	Unmatched	15.60	14.93	0.67	0.24	2.82***
	ATT	15.60	16.12	-0.52	0.36	-1.46
Size, Log Square	Unmatched	11.77	11.25	0.52	0.31	1.7
	ATT	11.77	12.44	-0.66	0.45	-1.46
Asia	Unmatched	0.19	0.31	-0.12	0.05	-2.37**
	ATT	0.19	0.23	-0.04	0.07	-0.6

Europe	Unmatched	0.18	0.02	0.16	0.02	8.33***
	ATT	0.18	0.12	0.06	0.06	1
Africa	Unmatched	0.44	0.43	0.01	0.06	0.25
	ATT	0.44	0.48	-0.03	0.09	-0.37
North America	Unmatched	0.07	0.14	-0.08	0.04	-2.05**
	ATT	0.07	0.09	-0.02	0.05	-0.41
South America	Unmatched	0.06	0.04	0.02	0.02	0.9
	ATT	0.06	0.03	0.02	0.04	0.58
Alliance of Small Island	Unmatched	0.17	0.27	-0.11	0.05	-2.17**
	ATT	0.17	0.11	0.06	0.06	0.9
Former Warsaw Pact	Unmatched	0.11	0.01	0.10	0.01	7.41***
	ATT	0.11	0.01	0.10	0.04	2.66***
British Commonwealth	Unmatched	0.40	0.46	-0.06	0.06	-1.09
	ATT	0.40	0.44	-0.04	0.09	-0.5
Francophonie	Unmatched	0.39	0.28	0.11	0.05	2.25**
	ATT	0.39	0.34	0.04	0.08	0.52
Island	Unmatched	0.21	0.26	-0.05	0.05	-0.98
	ATT	0.21	0.18	0.03	0.07	0.48
Landlocked	Unmatched	0.23	0.21	0.03	0.05	0.55
	ATT	0.23	0.19	0.04	0.08	0.58
Federal	Unmatched	0.08	0.11	-0.03	0.03	-0.97
	ATT	0.08	0.13	-0.06	0.06	-0.89
Year	**Unmatched**	**1996.98**	**1994.46**	**2.52**	**0.70**	**3.6*** **
	ATT	**1996.98**	**1992.39**	**4.59**	**1.13**	**4.06*** **
Propensity Score	**Unmatched**	**8.07**	**6.31**	**1.76**	**0.80**	**2.22** **
	ATT	**8.07**	**8.05**	**0.02**	**0.17**	**0.14**

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

Total Observations: 863

Untreated Observations: 773

Treated Observations: 90

Table I-4
Summary of Matching Results, Single Nearest Neighbor Matching

	Investment	GDP per Capita Growth	Trade	Inflation	Government Consumption
7 Years Pre-Adoption	0.14	0.03	0.87	-0.32	0.86
6 Years Pre-Adoption	-0.52	0.51	0.48	-0.77	0.67
5 Years Pre-Adoption	-0.36	-1.22	0.52	-0.36	0.79
4 Years Pre-Adoption	0.41	-1.45	0.83	0.35	0.99
3 Years Pre-Adoption	-0.27	0.32	0.95	0.25	0.83
2 Years Pre-Adoption	-0.55	0.67	1.07	0.48	0.92
1 Year Pre-Adoption	0.00	0.59	0.86	0.00	0.57
Year of Adoption	-0.37	2.17**	0.88	0.40	0.18
1 Year Post-Adoption	0.31	0.38	1.10	-0.68	0.58
2 Years Post-Adoption	0.69	1.20	1.09	-0.60	0.47
3 Years Post-Adoption	1.11	1.79*	1.16	-1.51	0.65
4 Years Post-Adoption	0.97	1.93*	0.80	-1.82*	0.35
5 Years Post-Adoption	1.21	1.38	0.87	-1.89*	0.01
Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**1.69***	**0.49**	**-0.25**	**-1.97***	**-1.75***
Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**-0.79**	**0.37**	**0.35**	**0.47**	**-0.21**
Treated Observations	77	78	76	78	80
Untreated Observations	639	683	662	684	716

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

***Notes*: T statistics above for mean differences, treated (adopters) minus untreated (matched controls); Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-5
Comparison of Matching Methods

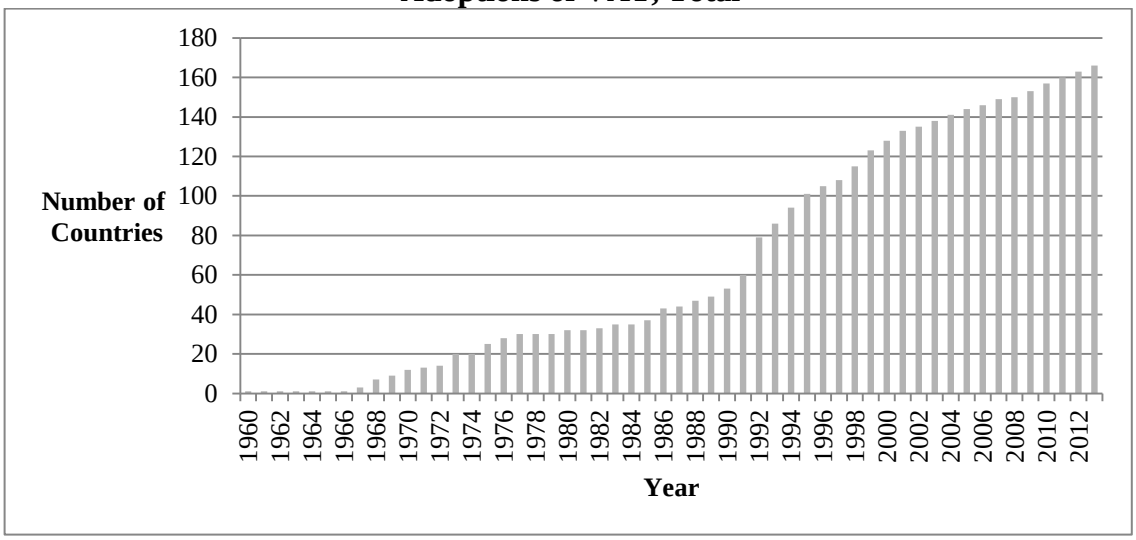
Variable, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	1 Nearest Neighbor, No Replacement	3 Nearest Neighbors	Radius, .025	Radius, .05	Kernel, Epanechnikov	Kernel, Normal	Total Significant at 10% level
Investment	2.13**	2.03**	1.56	1.60	2.40**	2.13**	5
GDP per Capita Growth	1.71*	0.64	0.22	0.79	0.06	0.85	1
Trade	0.23	0.47	-0.40	-0.25	1.32	0.94	0
Inflation	-1.41	-1.14	-0.94	-2.52***	-1.66*	-1.78*	4
Government Consumption	-1.77*	-1.53	-1.90*	-1.96*	-1.18	-1.14	4

Variable, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	1 Nearest Neighbor, No Replacement	3 Nearest Neighbors	Radius, .025	Radius, .05	Kernel, Epanechnikov	Kernel, Normal	Total Significant at 10% level
Investment	-0.74	-0.79	-1.20	-1.18	-1.28	-0.75	0
GDP per Capita Growth	-0.07	0.07	-0.03	-0.20	-0.30	0.11	0
Trade	-0.02	0.00	-0.05	-0.16	0.01	0.20	0
Inflation	0.08	0.29	1.33	1.49	1.01	0.05	0
Government Consumption	-0.46	-0.30	-0.49	-0.31	-0.41	-0.34	0

*** $p < .1$, ** $p < .05$, *** $p < .01$**

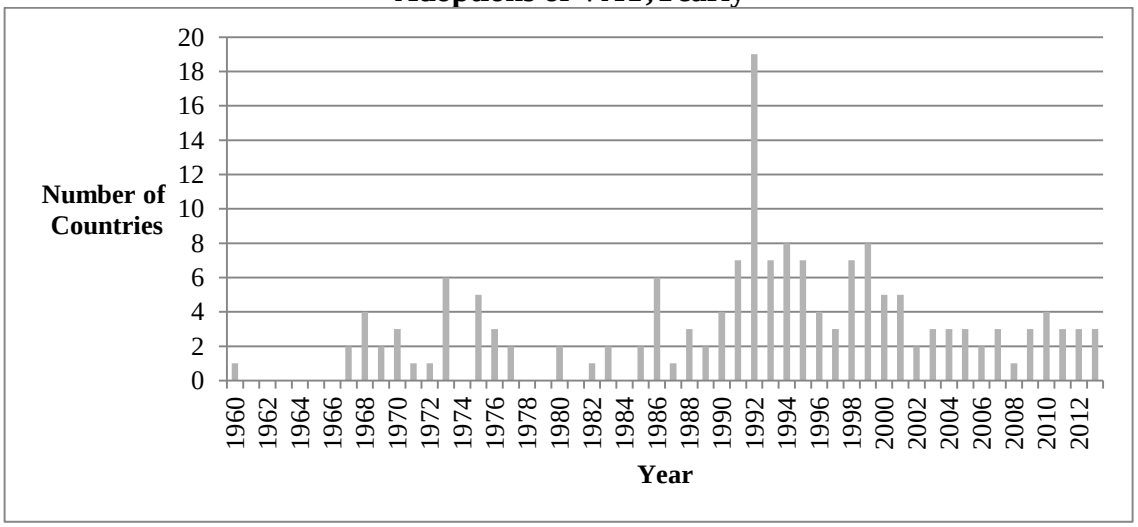
***Notes*: T statistics above for mean differences, treated (adopters) minus untreated (matched controls); Propensity scores for matching are generated by hazard regression (1) on Table I-2. Total significant also counts results from the single nearest neighbor on Table I-4.**

Figure I-1
Adoptions of VAT, Total



Source: Author's Calculations

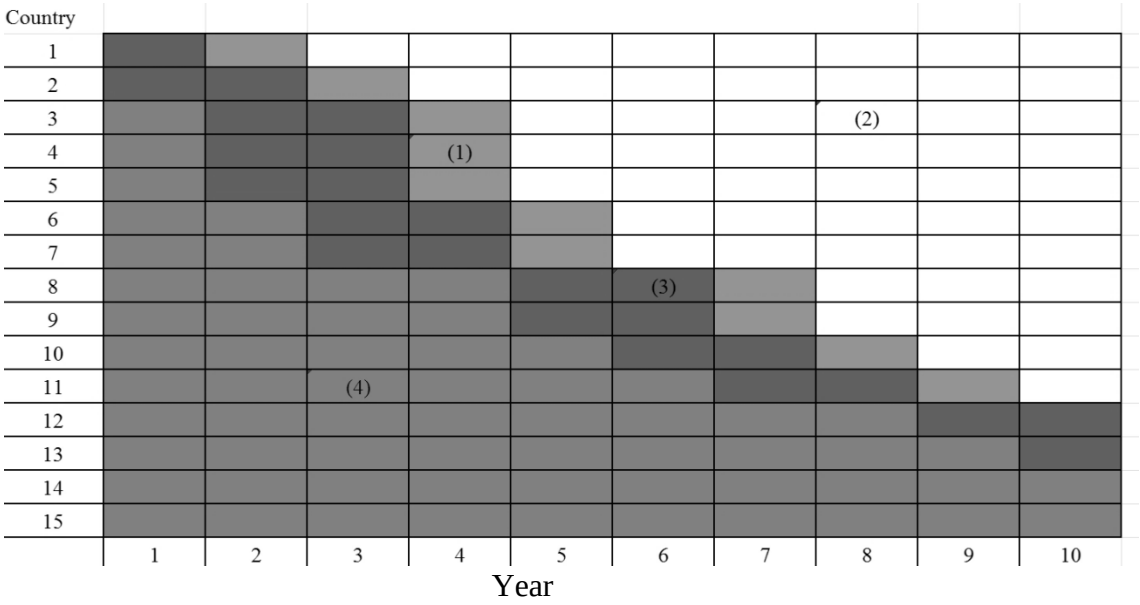
Figure I-2
Adoptions of VAT,Yearly



Source: Author's Calculations

Note for Figures I-1 and I-2: Vietnam in the 1970s, Grenada in 1986, Ghana in 1995, Malta in 1998, and Belize in 1999 all passed a VAT and later repealed it. Vietnam, Grenada, and Ghana would later re-adopt a VAT in 1999, 2010, and 1998 respectively, and these re-introduction events are included in this paper. Belize and Malta are included with original adoption dates. This is in line with what was done by Keen and Lockwood (2010). Omitting these observations does not qualitatively change results.

Figure I-3
Visual Explanation of Matching



Region (1) are country-years that have adoption events. These are the treatment observations. Information from these country-years are used in both the selection equation and matching.

Region (2) are country-years after adoption events. Information from this region is neither included in the selection equation nor the matching equation.

Region (3) are the 5 country-years pre-adoption that cannot be used for matching due to the need to have control country-years that do not adopt VATs for 5 years after being assigned as a control. They still provide information to the selection equation.

Region (4) are country-years 5 years or more before an adoption event. This is the pool of controls. Information from these country-year pairs are included in both the selection equation and matching.

Matching for this paper covers more countries than the 15 listed and a longer period than the 10 listed above, and so this figure is presented solely for demonstration purposes.

II. The Effect of VATs on Government Balance Sheets

1. Introduction

Over the past half century the value added tax, or VAT, has become a major source of government funding for a significant number of countries around the world. Hines and Summers (2009) argue that this trend will continue, with consumption taxes like the VAT becoming increasingly important due to the increasing mobility of economic activity that will make it harder to assess income taxes. Countries often consider passing a VAT in order to solve federal budgetary shortfalls, considering it a particularly effective strategy to raise revenue and eliminate deficits as it is able to collect more revenue for a given amount of government effort than most other taxes (Keen 2010). Similarly, Ebeke and Ehrhart (2012) have found that countries with VATs tend to have lower variance in tax revenues as a share of GDP relative to countries without this tax, making the VAT a useful tool at revenue smoothing for developing countries. Widespread belief in the VAT's ability to gather revenue, solve deficits and reduce variance in tax receipts is apparent recently in the words and actions of lawmakers, as well as those of international entities such as the IMF. Theoretical papers agree with these findings, with Diamond and Zodrow (2013) considering the effectiveness of a VAT in reducing US government deficits in a CGE model. They find the tax it is able to rectify government budget issues but remain relatively progressive, as it is able to tax a broader base with fewer distortions and less enforcement effort. However, it is not clear that the increase in government revenue will not then be immediately offset by increases in government spending due to lower cost of collecting funds, and so the net effect of the VAT on debt remains ambiguous.

This paper addresses the ability of a VAT to improve government balance sheets. I specifically test the VAT's effects on various measures of debt, government deficits, tax revenues and public expenditures as a share of GDP in a broad sample of countries that adopted the VAT between 1986 and 2010[18]. The vast majority of countries in this sample are developing countries, as developed countries adopted the VAT before this period. I consider both the level of these variables, as well as their variance, as VATs may improve the stability of some of these fiscal indicators in addition to affecting their volume. VATs are likely to affect government revenue, government spending, and GDP growth simultaneously, so the policy's impacts on all of these measures are not immediately clear.

While previous studies of the VAT use regression analysis, I follow Ufier (2014) in employing a matching framework to evaluate the fiscal impact of VATs. This method helps to address estimation problems in this setting. As a country chooses whether to adopt a VAT, adoption is a choice and naïve OLS estimates will not be capture the effect of the tax alone, but also pick up the effect of selection into the tax. Not all countries without a VAT are relevant comparisons to countries who adopt the VAT as countries select the treatment endogenously. Using matching should alleviate the issue of selection on observables by comparing countries that adopted the VAT only to those countries which are similar, but did not adopt the tax.

Estimation in this paper will follow a two-step process. First, using survival analysis, I will estimate the probability of a country adopting a VAT in a given year

[18] While I model adoptions in years following 2007, I cannot look at the impact of these adoptions as I lack sufficient data on government balance sheets to look at the impact of taxes in these most recent years of adoption.

based on observable factors. Second, I will match two (or more) countries with similar probabilities of failure[19], in which one adopted the VAT and the others did not, in order to estimate the effect of a VAT in the 5 year span following its adoption. As a falsification test, I will also compare adopters and similar non-adopters for 7 years before adoption in order to verify there were no pre-existing differences or trends that would invalidate the assumption countries were similar prior to choosing the treatment.

Despite its importance and promised benefits, the VAT is relatively understudied in economics. Several such studies including Keen and Lockwood (2010), Nellor (1987) and Ebeke (2010) look at the determinants of the VAT and generally find it increases tax revenues. Stockfisch (1985) considers its impact on government size in OECD countries, finding both adopter and non-adopter governments grow over time, with no differential effect of growth coming from the VAT. Ufier (2014) examines its macroeconomic impacts, finding it to be generally positive for investment and growth. Yet concerns still remain about the longer term fiscal impact of such a tax.

Some policy makers fear a "money machine" with the VAT leading to the government's spending of all the additional funds it takes in, as it is easier to collect taxes and they thus increase collection efforts and immediately spending them, leaving balance sheets no better off than before the passage of the tax (Keen 2007). In previous general tax papers, authors had found that higher tax receipts did lead to higher tax expenditures in some cases (Manage and Marlow 1986) but not in others (Von Furstenberg et al 1986.) Lee, Kim, and Borcherding (2013), however, propose the causation flows the other direction- instead of high revenues causing higher demand for

[19] A "Failure" in this case is referring to VAT adoption.

public goods, people's preferences for higher levels of public goods tends to lead them to pass more revenue collecting measures. The opinion on the VAT is thus mixed, and more solid empirical evidence of its impact on debt and deficits would help to better understand the possible policy uses of a VAT.

This paper is organized as follows. Section 1 provides a brief summary and introduction, and section 2 provides background on the VAT. Section 3 discusses the first part of the estimation, survival analysis, and section 4 discusses the second, matching. Section 5 covers results and section 6 concludes.

2. Background

France passed the first general VAT in 1954, reforming it significantly in 1968. By this time, Cote d'Ivoire had already passed a VAT in 1960 and the tax was beginning to spread to Latin America and Western Europe. Initial adoptions came slowly in the 1970s and early 1980s, spreading eventually to Southern Europe and Sub-Saharan Africa. By the early 1990s many countries in Eastern Europe and Asia began to enact the tax. In the 2000s, small island nations in the Pacific and Caribbean became one of the last major groups to implement the tax, and currently the Middle East remains the only major holdout region, with the Arab Gulf states currently investigating a VAT in the future. The sole major industrialized economy without a VAT as of 2015 is the United States.

2A. Further Literature

A VAT, like any other tax, can be used to solve budget deficits and improve government balance sheets. However, it is not at all clear that the tax would be used to actually correct budget sheet imbalances faced by countries adopting it. The static benefits of raising revenue through an easy to collect tax mechanism may be counterbalanced by the dynamic effects of governments instead choosing to raise expenditures or cut other taxes, leaving its fiscal position no better than before the tax (Keen 2007, Toder and Rosenberg 2010). Additionally, Bird and Smart (2012) highlight the increasingly prominent role of the VAT as a tax intended to shore up large entitlement programs whose original funding streams prove insufficient. Theoretical papers may predict the impact of a VAT or other consumption tax on macroeconomic variables, but often lack empirical tests of their theories. While Ufier (2014) considers the impact of a VAT on government consumption as a share of GDP, and Ebeke and Ehrhart (2012) considers stability of revenues in a regression framework, empirical evidence does not definitively state whether VATs tend to reduce debts and deficits in countries that have passed them.

Numerous nations in recent years are looking at a VAT for the reasons of raising revenue and reducing deficits, usually under advisement from the IMF or in official IMF reports. Bird (2010) claims that the IMF often promotes a VAT to countries which are experiencing high budget deficits. Aruba's recent fiscal reforms selected to replace many current taxes with the VAT (IMF 2010). The Bahamas found numerous tax deficiencies and suggested reforming the government's tax structure by imposing a VAT, as it does often with many recent article IV missions (IMF 2013.) The new

government of Afghanistan is considering the VAT along with taxes on natural resources to be its primary sources of revenue (IMF 2012). Burundi's stated goals in passing the VAT is to bring its tax code closer in line with the international standard, their de facto endorsement of the VAT as the standard world tax (African Development Bank 2010). The Marshall Islands considers a VAT likely to be a successful means of securing revenue for the economy, if accompanied with other fiscal reforms (IMF 2008). Prior to acquiring the VAT, Palau considered their tax system in gross need of modernization and reform, with the VAT being the best way to reform it for fiscal stability (IMF 2012). More directly, the IMF recently demanded implementation of the VAT in Pakistan with no exemptions or changes allowed; If Pakistan were to try to alter the law as presented to them, they would cease future assistance in the country (Rana 2013). There are thus a fair number of countries considering a number of tax reforms recently, and their general consensus has been to side with the VAT as the best available of those options for raising revenue and keeping it stable.

The last major bloc of countries without a VAT, oil producing countries in the middle east, are now considering the VAT for revenue stability purposes. The Gulf Cooperation Countries, long reliant on oil revenues and banking taxes for funding, consider this tax to be the best means of acquiring a constant stream of revenue and reversing government deficits (Charalambous 2012). Officials in Bahrain, Kuwait, UAE, and Qatar have all spoken out in support of this action (Izzak and Saleh 2013, Kapur 2012, Peninsula 2012, Rafique 2011.) There is a great deal of international support behind the VAT as a tool for solving budget deficits or raising revenue in the

long run. Yet for all the strength with which this tax is promoted, there is a relative dearth of empirical evidence.

3. Survival Analysis

In order to properly account for the selection in acquiring to VAT, one first needs to model the decision to adopt a VAT. This paper uses the same selection equation and many of the same covariates to model VAT adoption as Ufier (2014), employing a Cox proportional hazard model with countries surviving in a VAT-free state for a period until they acquire the tax and thus exit the estimation. I will only briefly cover the topic of selection equation and its determinant in this paper, as Ufier (2014) covers it in greater detail. This paper uses a propensity score matching technique, modeling the adoption of a VAT as a Cox proportional hazard model. The persistence of a VAT means that a survival model would provide a good fit to the observed data, as countries tend to acquire the tax and never repeal it. This analysis allows me to estimate the probability of failure, adopting a VAT, in each country in each period. I then compare countries at equal risk of failing in the case where one failed and the other did not in order to obtain an unbiased estimate of the effects of adopting a VAT, as actual time to failure is random conditional on countries having the same probability of failure (Lu 2005).

3A. Determinants of Adoption

There is a great deal of theoretical guidance as to the determinants of a country choosing to pass a VAT, but the major determinants of tax adoption can be split into

several groups. International pressure, in the form of IMF lending, could lead to more countries acquiring the VAT (Bird 2010). IMF lending is represented here as a dummy variable equaling 1 if a country is under an IMF lending program at any point in a given year. Proximity to other countries with a VAT could also lead to one acquiring it, either from the experience of seeing a neighbor implement it or tax competition, and so the share of neighboring countries with a VAT is included as a determinant (Cizek et al 2012). National economic composition, specifically countries with a large share of GDP coming from agriculture or natural resource production, likely affects the type of taxes a government chooses to collect. Countries that collect revenue from natural resources may be less likely to rely on other types of taxes, and agriculture is considered difficult to monitor for providing reimbursement and collecting taxes necessary with a VAT system (Keen and Lockwood 2010, Cnossen 1998).

Geopolitical factors including landlocked status, island status and log land area are all included as dummy variables. These measures all potentially affect a nation's decision to acquire a particular tax due to their effects on enforcement, with landlocked countries having more difficulties in controlling smuggling of products to avoid VATs compared to countries with only a limited number of points of entry (Keen 2010). Historical factors, including French or British influence, as well as Warsaw pact membership all could affect a countries' legal system and thus decision to adopt the tax (Keen 2010). Decentralized federations may also avoid a tax that requires centralized collection like a VAT (Ebrill et al. 2001, Treisman 2002). As VATs require a great deal of institutional capacity to implement, I include the Polity IV measure of democracy as

a measure of institutional quality due to its relatively good coverage over time and across countries (used by authors such as Ang 2010.)

Demographic factors may affect the optimal tax system for a country, especially for ones with large safety nets for the elderly (Bird and Smart 2012). A tax system with a fixed setup cost may also not be considered worthwhile in very small or very poor, resource constrained countries. Therefore, I also include the log of population, percent of population above 65, and log of GDP per capita as determinants of adoption. Next, I include the macroeconomic outcome variables of imports, exports, inflation[20], investment and government consumption all as a share of GDP. These variables may determine legislator's decisions to implement the tax.

Finally, as there is usually some lag between the planning and passage of a VAT and the tax actually being collected, I lag all determinants by 2 years. Thus, a country choosing to acquire a VAT in 1996 does so based on values of its covariates from 1994. This is in line with Ebrill et al. (2001) who state there is roughly an 18 month lag between and decision to pass the VAT and its collection, as well as Cizek et al (2012) and Ufier (2014) who also use a 2 year lag. I also include regional dummies for Africa, Asia, Europe, North America, and South America, with the pacific being the omitted category.

3B. Summary Statistics and Results

Table II-1 presents summary statistics on determinants of VAT adoption, as well as the outcome variables affected by the VAT policy. The independent variables used in

[20] I use the natural log of 100+the rate of inflation in an effort to smooth relatively frequent extremely high inflation events among countries.

the survival analysis have coverage from 1986-2010, allowing me to model countries' VAT adoptions during this period. I consider seven outcome variables that summarize government fiscal positions, coming from three different data sets. First, the Milesi-Feretti and Lane (2011) external debt data set covers essentially the entire period of interest, combining externally held private as well as public debt. Second, the World Development Indicators also collects an external debt (as a share of GDP) series, and it covers approximately 80% of the country-years during the period of interest. Finally, the World Economic Outlook contributes five data series to this study: central government debt, deficit, cyclically adjusted deficit, tax revenue, and government expenditure, all as a share of GDP. These series provides numerous alternative ways to consider the fiscal impact of a VAT, although it has smaller coverage than either of the above measures.[21] Any null findings using this debt data may be due to the low power coming from a smaller sample size rather than evidence of a lack of robustness across data sets.

Table II-2 shows the result of the Cox proportional hazard regression used to model the decision to adopt a VAT. In each year, the country has a probability of either surviving in a VAT-free state or adopting a VAT and exiting estimation. The probability of adopting a VAT in each period is affected by 25 variables, both time variant and invariant, described in section 3.1. Seven variables are found to be statistically significant determinants of choosing to enact a VAT. Countries receiving IMF assistance are more likely to pass a VAT as well and countries with a large number

[21] The World Development indicators have data on many of these measures as well, but they cover less than 10% of the country years being studied and so may not provide useful results.

of neighbors that already have the VAT. This would be consistent with the idea that VATs spread by proximity and that IMF pressure plays a role in adoption. Countries with larger populations or higher GDPs per capita also are more likely to get the VAT, which is consistent with the idea that one may be better able to spread the fixed cost of the tax in larger countries, and that wealthier countries have an easier time administering it. Countries with large natural resource rents as a share of GDP are less likely to get the tax, being able to rely on other sources of funding. Countries governed as federations are less likely to pass a VAT potentially due to difficulties in administration. However, British Commonwealth countries are also less likely to pass it. These are all generally consistent with the previous literature.

4. Matching

I use the results of the Cox proportional hazard model selection equation from Table II-2 to generate propensity scores. For each country-year pair, I save the instantaneous probability of adopting a VAT from the hazard model. Using the process outlined by Lu (2005), for two observations wherein the difference in failure rates approaches 0, failure is random conditional on those observables. Thus I will be comparing country-years that fail and acquire VATs to country-years that do not fail but were otherwise similar.

Several methods are available to produce matching estimates. Most popular is nearest neighbor matching. I can compare the n nearest neighbors to a single treated observation, comparing the treated units to the average outcome of the n closest untreated country-years, both with or without replacing the pool of untreated units after

performing a match. Alternatively, one could compare treated units to all untreated units with a propensity score within a given distance- a propensity score distance of .05, for example- of the treated unit. Finally, one could compare the treated unit to all untreated units, but weight them inversely based on the distance between the treated and untreated units propensity score. Thus, those with very similar propensity scores make up a larger share of the comparison group than those with very different ones. While one could use linear weighting, there are a number of other weighting kernels one could employ for this matching. This paper primarily uses the inverse propensity score weighting method employing a normal kernel, or normal kernel matching for short. I employ several other methods as a robustness check to see if results vary by matching methodology.

Figure I-3 (from the first chapter of this dissertation) presents a diagram for illustrative purposes of country-years that will be studied by this matching method. Going from time 0 to time T, countries acquire a VAT, and region (1) of these country-years, forming the treatment group. Country-years after this point already have VATs, in region (2), and are not used for matches and do not contribute to survival analysis. As I will be considering the effect of a VAT 5 years after its adoption, I cannot use any country-year observations less than 5 years before that country adopts as a control, in region (3). If I do, some control units would acquire the treatment during the studied 5-year period and bias the treatment effect of the VAT downward. They still can be used to estimate VAT adoption probabilities. If a country will not get a VAT for 5 or more years, and has not yet adopted one, region (4), it can be used a control observation for matching as well as affect the survival equation. Therefore I compare treated country-years from region (1) to untreated country years from region (4) to obtain an unbiased

estimate of the effects of adoption of a VAT. There is no time restriction on matching, so adoption events in a given country-year can be matched to a different country that is also in a different year. Restricting the time dimension to match only within the same year will significantly reduce the number of control observations available for a given adoption event, especially in later years when many countries have adopted a VAT and thus exited estimation.

4A. Identification Assumptions

In order for the estimation method to be valid, the matched controls need to be similar to the treated units on all observable covariates. If the two groups differ on average on any covariates, it will not be clear if the difference in outcomes among groups comes from the treatment or one of the covariates. Table II-3 summarizes these differences. Prior to matching, the treated and untreated groups differ on 11 of the 25 covariates at least at a 10% level. After matching, they differ on zero covariates at least at a 10% level. Further, Figure II-1 provides the distribution of propensity scores of adopters versus non-adopters. There is a large area in which there is overlap of propensity scores, or common support, where adopters and non-adopters both have similar probability of adopting the treatment, which is necessary for this method to function. The model also correctly assigned the adopters at higher risk of adopting the tax than non-adopters, as seen by the difference in peaks of the density function. Table II-3, marked in bold at the bottom, shows that prior to matching, the propensity scores of adopters versus non-adopters differed by 10 standard deviations, indicating a large difference in initial risk of adoption. After matching, they differ by only .14 standard

deviations, indicating that adopters and non-adopters they are at similar risk of adoption based on observables.

Additionally, while not explicitly a covariate and hence marked in bold on the table, VAT adoption country-years tend to be later in the panel than either the matched or unmatched controls. As countries exit estimation over time when they adopt a VAT, there should be more country-years available for controls at early dates then at later ones as seen in Figure I-3 (from the first chapter of this dissertation). While the difference in years could be cause for concern, possibly affecting outcome variables, it would have to affect the observed outcome variables through a channel other than one of the included covariates, as there is no statistically significant difference between the groups in all but one of those variables. This evidence indicates that the matching method significantly reduces the bias resulting from a country's selection of tax policy, allowing me to estimate the treatment effect of a VAT more accurately.

5. Results

Using the failure rates generated in Table II-2 as propensity scores, one can compare countries with equal probabilities of adopting a VAT in cases where one country did and the other countries did not. Countries are matched using a normal kernel, in which countries with propensity scores closer to the adopter are weighted more heavily than those further from the adopter. As demonstrated in Table II-3, this matching significantly reduces the selection on the observables, leaving the control and treatment groups on average nearly identical on observable variables. Comparing their mean outcomes should reduce the bias of the estimate of the treatment effect of passing

a VAT, as one is comparing only similar countries, making the control countries a relevant comparison to the treatment group. Table II-4 compares the treated units to matched controls.

5A. Central government Debt

There is reasonably strong evidence of reduction in central government debt from the WEO data series as indicated by (1) on Table II-4. Going from the period two years before adoption to five years after, the average debt levels fell in treated countries compared to untreated ones as indicated by the negative and significant T statistic.[22] This corresponds to an average fall in debt as a share of GDP of 41.23 percentage points. The fact this sample is relatively small would mean low power, making a significant effect here all the more surprising. However, there is no evidence of changes in variance in debt after adoption, from the two years prior to five years after. There is no systematic change in debt variance or levels prior to adoption of treatment (seven years to two years before), which one would expect to see if there were no pre-existing difference in trends.

5B. Deficits

There is evidence of changes in Deficits from the WEO data series (2) as well as primary adjusted deficits (3) in the years following the VAT adoption. In the span of time after adoption, the change from two years before adoption to five years after, average deficits fell (or surpluses rose) in treated countries compared to untreated ones as indicated by a positive T statistic significant at a 5% significant level. While the

[22] Recall the determinants are lagged 2 years in the selection equation.

effects are also positive for primary deficits, the T statistic is not significant, potentially due to the larger smaller sample size. This corresponds to approximately 3.63 percentage points of GDP, and thus is evidence the VAT can contribute to reducing deficits. There is no evidence of such a change over time in the variance of surplus/deficits following the passage of a VAT. However, one point of concern is that while there was no difference in levels prior to adoption (seven to two years before), deficits worsened in adopters relative to non-adopters in the pre-period by 2.89 percentage points of GDP, indicating that the surpluses afterward are roughly cancelled out by the deficits prior to getting the tax.

5C. Revenue and Expenditure

Next, there is also a statistically significant effect of VATs expenditures (4), but not on tax revenue (5), relative to comparison countries. Expenditures fall significantly at the 1% level in the countries adopting the tax, dropping by 5.09 percentage points of GDP relative to comparison countries from the period two years before adoption to five years aftr. This would argue strongly against the money machine theory, and combined with the evidence of falling government debts, indicate that countries took the time around VAT adoption as an opportunity to improve their fiscal balance sheets as a whole. Revenues on the other hand, show no statistically significant effect, doing worse than comparison countries, with their revenue increasing by 1.94 percentage points of GDP while comparison countries see a gain of 3.42 percentage points of GDP in the same period. This is significant evidence of tax substitution behavior, with the VAT replacing other taxes as sources of revenue, rather than being used to increase overall

revenue. There is no evidence of changes in variance of either measure after, nor is there evidence of pre-existing differences in levels or trends of either revenue or expenditure, or their variances.

5D. Externally Held Debt

Neither data set, World Development Indicators (6) and Milesi-Feretti (7), show a statistically significant change in debt in the years following VAT adoption, looking from 2 years before adoption to 5 years after adoption. However, as this is a combined measure of publically and privately held debt, and there is a fall in government debt in some of our measures, this may be indicative of changes in behavior among private individuals.

5E. Comparison to Results without Matching

To show how results would differ without correcting for selection, Table II-5 compares the difference between VAT adopter country-years and non-adopter country-years before any matching is applied. Thus any differences will be capturing not only the treatment but also the selection into the treatment, including countries that are not relevant comparisons to adopters, which may be biasing the results.

Adopters, when compared to all non-adopters appear to have significantly higher central government debts (1), which fall over time both prior to and after getting the VAT. Matching allows one to narrow this down to only seeing changes in debt after getting the VAT, so that the drop in debt may now plausibly come from the tax instead of pre-existing differences. Additionally, variance is much higher for adopters both

before and after adoption, which does not appear in the matched results. This would indicate that the observed higher variance in debt experienced by VAT adopters comes from comparing it to inappropriate comparison countries.

Findings on deficits (2,3)in Table II-5 before matching do not differ too much from the after matching results on Table II-4. Adopters tend to run lower deficits compared to all other non-adopters following getting the tax, but run higher ones prior to getting the tax. However, unmatched results do show a lower variance in deficits for adopters than non-adopters, which disappears when better comparisons are used in matching in Table II-4.

Without using matching and accounting for selection on observables, countries that eventually acquire a VAT have lower expenditures and tax revenue as a share of GDP (4,5) in the years prior to as well as after adoption. Variance in both falls after adoption, however. Note this agrees with the findings of Ebeke and Ehrhart (2012), which found that VAT reduced volatility in tax revenue. However, there were no such differences in countries when using the matching methods, and thus these differences may be driven by selection into the treatment rather than the treatment itself.

VAT adopters do not have higher external debt loads (6,7), albeit the world development indicators show a slight drop in debt in the fourth and fifth years post adoption. This agrees with the mostly null results found in the matching section.

These effects, taken together, might lead one to infer that the VAT will reduce variance in government expenditure, tax revenues, and surplus/deficits, as well as lower government debt and improve deficits as a share of GDP, a significant improvement in stability of government operation and government balance sheets. However, these

differences in means are not capturing the effect of VAT adoption alone, as one is comparing countries that acquire the VAT to all other countries. As countries passing the tax likely look different on a number of observables as well as un-observables, this difference represents not only the effect of the treatment but also the selection into the treatment. After controlling for selection on observables, there is still evidence instead for an improvement in central government debt and deficits, but no evidence for changes in variance of any of the other measures. However, matching does detect a general fall in government expenditures that the unmatched results do not show. Matching methods thus show the fears of VAT as a money machine are unfounded, while showing it still has a robust effect on debt and deficits. On the other hand, matching methods also find little evidence of VAT's revenue stabilizing effects.

5F. Alternative Matching Methods

The findings in Table II-4 could be driven by the choice of matching method, and so several different matching methods are attempted in Table II-6. To conserve space, only overall changes in outcomes after VAT and before VAT are compared. No matching methods should show statistically significant differences changes in the years prior to VAT adoption or else the validity of the method may be called into question, while if the VAT has an effect that there should be some change in the years after the tax is passed. The additional methods used in this case include single nearest neighbor with and without replacement, three nearest neighbor with replacement, caliper matching of size .1 and .05, as well as Epanechnikov kernel weighting to accompany normal kernel weighting.

Only several of the results are robust to alternative matching methods. All matching methods employed show a decline in central government debt as a share of GDP following adoption of a VAT, with no pre-existing difference before adoption. Both measures of surplus and deficit, while showing improvements in surpluses following adoption of the tax, also saw them worsen beforehand, indicating a difference in trends between adopters and non-adopters prior to getting the tax. These results may thus not be valid. There is consistently no effect for changes in tax revenue as a share of GDP, while expenditure falls in all cases, with generally no evidence of pre-existing differences in trends. There is some weak evidence of falls in externally held debt in the world development indicators series, but not in the Milesi-Feretti series.

Table II-7 compares the changes in variance of outcomes before and after the VAT treatment under various matching methods, finding a relatively null picture. Two notable events are that government expenditures among adopters, while seeing no effect under kernel matching, have statistically significantly lower variance under all other matching methods. This would be evidence that a VAT reduces variation in government outlays at the same time it makes them smoother. Secondly, there is significant evidence of a pre-existing difference in variances of externally held debt by the World development indicator series among VAT adopters, with countries that adopt the tax having lower variation in debt levels prior to adopting.

After performing this robustness analysis, several clear results emerge. First, countries that adopt the VAT reduce their debt as a share of GDP, indicating they do get their fiscal balance sheets in order. Second, there is no evidence of a money machine effect of the VAT, as countries getting the tax see their expenditures fall and become

smoother. Lastly, there is a great deal of tax substitution behavior with countries replacing other taxes with the VAT, as there is not a large increase in tax revenues as a share of GDP. These findings are all potentially affected by changes in GDP affecting the denominator of these measures, as Ufier (2014) did find higher growth under a VAT.

6. Conclusion

The VAT has been considered or implemented in numerous countries as an important tool for raising revenue in order to reduce government deficits and debts, being entertained recently even by policy makers in the largest VAT-free nation, the United States, as well as many developing countries Yet there is mixed opinion as to whether this tax is actually effective at lowering deficits and decreasing debt. This paper has provided novel evidence that a VAT can indeed be an effective tool in reducing debt while at the same time they do not lead to a large increase in government expenditure. Without employing matching methods, I would have not detected these effects. This study found less evidence of the effect of VAT on revenue levels or stability.

While the findings of this paper are suggestive and may push many holdouts to once again consider acquiring a VAT, the results should be approached with caution. First, the method does not differentiate between various other factors that are important to the success of a VAT, including taxation rates and operation of the rebate system (Edmiston and Fox 2006). Popular VAT exemptions given to staples, such as food, may be expensive and ultimately unwise policies (Iorwerth and Whalley 2002.) Second, this

paper does not explicitly consider the welfare effects of the tax or its incidence of payment, both of which could be of concern to policy makers due to the burden placed on consumers and poorer individuals (Politi and Mattos 2011, Stiglitz 2008, Emran and Stiglitz 2005.) Third, this analysis is only able to address countries that have gotten VATs from 1986-2010, thus only studying some developing countries that implemented a VAT and fewer developed ones.

Finally, this paper uses numerous data sets in order to show the proposed effects on debts and deficits, and many of these data sets are missing a significant number of observations. Assuming their omission is random, this is not cause for concern. If selection is at work, it could be biasing the effects of a VAT upward, as countries missing data may have poorer institutions that are less able to benefit from the tax or implement it. If the sample size were to increase the power of the test for difference in means should increase, likely strengthening the evidence of effects of a VAT. Yet despite some shortcomings, these findings are still highly supportive of the fiscal balancing effects of a VAT and should provide a basis for further study of the tax.

Table II-1
Summary Statistics, Right Hand Side Variables, 1986-2010

Variable	Obs.	Mean	Std. Dev.	Min	Max
Dependent					
WEO Central Government Debt	490	82.35	97.85	0	931.66
WEO Surplus or Deficit	675	-2.30	7.03	-46.24	31.81
WEO Primary Surplus or Deficit	283	1.13	7.07	-17.36	31.81
WEO Tax Revenue	682	26.55	11.59	4.49	68.37
WEO Expenditure	690	29.08	10.83	6.44	82.08
WDI External Debt	832	105.50	125.31	2.94	1380.77
MF External Debt	1030	88.79	128.43	2.85	1620.28
Independent					
Agriculture	1031	24.74	17.13	0.11	80.07
Bordering Countries with VAT	1031	0.29	0.34	0	1
British Commonwealth	1031	0.38	0.49	0	1
Exports	1031	34.01	26.48	2.52	184.96
Federal	1031	0.15	0.35	0	1
Francophonie	1031	0.34	0.47	0	1
GDP Per Capita	1031	7.08	1.49	4.74	10.97
Government Consumption	1031	13.80	10.47	1.11	69.83
IMF Lending Dummy	1031	0.61	0.49	0	1
Imports	1031	41.56	28.08	6.86	175.39
Inflation	1031	4.79	0.41	4.26	10.20
Institutions	1031	-1.93	6.58	-10	10
Investment	1031	21.66	10.27	-2.42	74.82
Island	1031	0.13	0.33	0	1
Landlocked	1031	0.27	0.44	0	1
Natural Resource Rents	1031	14.43	14.45	0	71.61
Population	1031	15.69	1.74	12.78	20.88
% of Population 65+	1031	3.93	2.33	0.43	14.62
Size, Log Square Kilometers	1031	12.18	2.04	6.51	16.05
Warsaw Pact	1031	0.03	0.16	0	1
Africa	1031	0.54	0.50	0	1
Asia	1031	0.32	0.46	0	1
Europe	1031	0.04	0.20	0	1
North America	1031	0.01	0.11	0	1
South America	1031	0.05	0.21	0	1

Notes: **Observational units are country-year pairs.**
Independent variables are lagged two periods

Table II-2
Cox Hazard Regressions for VAT Adoption

Variable	1986-2010
Agriculture	.008
	(.014)
Bordering Countries with VAT	1.067*
	(0.546)
British Commonwealth	-0.660*
	(0.353)
Exports	0.003
	(0.011)
Federal	-1.738***
	(0.526)
Francophonie	0.306
	(0.329)
GDP Per Capita	0.493*
	(0.282)
Government Consumption	-0.025
	(0.018)
IMF Lending Dummy	0.940**
	(0.428)
Imports	-0.007
	(0.012)
Inflation	-1.513*
	(0.811)
Institutions	0.040
	(0.026)
Investment	0.008
	(0.021)
Island	0.531
	(0.558)
Landlocked	0.058
	(0.320)
Natural Resource Rents	-0.049***
	(0.017)
Population	0.608***
	(0.182)
% of Population 65+	-0.090
	(0.107)
Size, Log Square Kilometers	-0.127
	(0.118)
Warsaw Pact	0.242
	(0.916)
N	1031
Pseudo-R^2	0.132
Log Likelihood	-243.354
Chi Squared	74.180
Subjects	92
Adoptions	75

Standard errors in parentheses
*** $p < .1$, ** $p < .05$, *** $p < .01$**

Notes: **Observational units are country-year pairs; Values are coefficients for linear hazards, not hazard ratios. A coefficient of 0 thus implies no effect. To find the hazard ratios, exponentiate the coefficient; All independent variables are lagged 2 periods; All estimates performed with regional dummies for North America, South America, Europe, Asia, Africa, and Pacific, the last being omitted for co-linearity.**

Table II-3
Balance of Covariates for Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Agriculture	Unmatched	24.91	23.88	1.03	2.13	0.48
	ATT	24.91	26.58	-1.67	3.65	-0.46
Bordering Countries with VAT	Unmatched	0.43	0.25	0.18	0.04	4.41***
	ATT	0.43	0.32	0.11	0.08	1.45
British Commonwealth	Unmatched	0.39	0.36	0.03	0.06	0.5
	ATT	0.39	0.49	-0.11	0.11	-1.01
Exports	Unmatched	31.58	36.00	-4.43	3.29	-1.34
	ATT	31.58	33.83	-2.25	5.59	-0.4
Federal	Unmatched	0.11	0.17	-0.06	0.04	-1.31
	ATT	0.11	0.13	-0.03	0.07	-0.39
Francophonie	Unmatched	0.41	0.29	0.13	0.06	2.29**
	ATT	0.41	0.45	-0.04	0.10	-0.38
GDP Per Capita, Log	Unmatched	7.08	7.20	-0.12	0.19	-0.63
	ATT	7.08	7.01	0.07	0.30	0.24
Government Consumption	Unmatched	11.73	14.63	-2.90	1.36	-2.14**
	ATT	11.73	11.61	0.12	1.94	0.06
IMF Lending Dummy	Unmatched	0.76	0.51	0.25	0.06	4.08***
	ATT	0.76	0.84	-0.08	0.09	-0.89
Imports	Unmatched	39.99	42.68	-2.69	3.45	-0.78
	ATT	39.99	41.19	-1.20	6.16	-0.19
Inflation	Unmatched	4.71	4.81	-0.09	0.06	-1.62
	ATT	4.71	4.70	0.01	0.03	0.46
Institutions	Unmatched	1.52	-3.50	5.02	0.76	6.59***
	ATT	1.52	0.57	0.95	1.47	0.64
Investment	Unmatched	21.96	21.90	0.05	1.33	0.04
	ATT	21.96	22.51	-0.55	1.94	-0.28
Island	Unmatched	0.15	0.11	0.03	0.04	0.82
	ATT	0.15	0.24	-0.09	0.07	-1.25
Landlocked	Unmatched	0.28	0.26	0.02	0.05	0.43
	ATT	0.28	0.29	-0.01	0.10	-0.13
Natural Resource Rents	Unmatched	9.64	16.84	-7.20	1.87	-3.85***
	ATT	9.64	7.87	1.77	2.27	0.78
Population, Log	Unmatched	16.01	15.54	0.46	0.21	2.16**
	ATT	16.01	16.00	0.01	0.41	0.03
% of Population 65+	Unmatched	5.08	3.53	1.54	0.26	5.96***
	ATT	5.08	4.97	0.11	0.59	0.19
Size, Log Square Kilometers	Unmatched	12.26	12.13	0.13	0.26	0.5
	ATT	12.26	12.01	0.26	0.44	0.58
Warsaw Pact	Unmatched	0.09	0.00	0.09	0.01	6.78***
	ATT	0.09	0.04	0.05	0.04	1.27
Africa	Unmatched	0.52	0.51	0.01	0.06	0.15
	ATT	0.52	0.51	0.01	0.11	0.12
Asia	Unmatched	0.21	0.38	-0.16	0.06	-2.79***
	ATT	0.21	0.24	-0.03	0.09	-0.29
Europe	Unmatched	0.13	0.01	0.12	0.02	7.03***
	ATT	0.13	0.15	-0.01	0.06	-0.23
North America	Unmatched	0.01	0.02	-0.01	0.02	-0.35
	ATT	0.01	0.00	0.01	0.01	1
South America	Unmatched	0.07	0.04	0.03	0.03	1.01
	ATT	0.07	0.07	0.00	0.06	0
Year	**Unmatched**	**1996.79**	**1994.9**	**1.88**	**0.83**	**2.27**
	ATT	**1996.79**	**1992.41**	**4.37**	**1.33**	**3.3**
Propensity Score	**Unmatched**	**3.55**	**1.94**	**1.61**	**0.16**	**10.08**
	ATT	**3.55**	**3.52**	**0.03**	**0.2**	**0.14**

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

Total Observations: 844
Untreated Observations: 784
Treated Observations: 60

Table II-4
Summary of Matching Results, Normal Kernel

	(1) WEO Central Government Debt	(2) WEO Surplus/Deficit	(3) WEO Primary Surplus/Deficit	(4) WEO Expenditure	(5) WEO Tax Revenue	(6) WDI External Debt	(7) MF External Debt
7 Years Pre-Adoption	0.93	1.16	0.91	0.12	0.59	-0.81	-0.06
6 Years Pre-Adoption	0.82	1.03	0.59	0.19	0.63	-0.8	-0.15
5 Years Pre-Adoption	0.53	0.19	0.51	0.41	0.41	-0.49	-0.05
4 Years Pre-Adoption	0.49	0.25	0.3	0.15	0.23	-0.25	0.05
3 Years Pre-Adoption	0.65	0.38	0.32	0.05	0.24	0.07	0.13
2 Years Pre-Adoption	0.65	-0.82	-1.01	0.82	0.24	0.2	0.17
1 Year Pre-Adoption	0.24	-0.18	-0.95	0.69	0.45	0.07	0.14
Year of Adoption	-0.47	-0.7	-1.02	0.84	0.23	-0.5	0.03
1 Year Post-Adoption	-0.34	-0.32	-0.66	0.57	0.18	-0.55	0.05
2 Years Post-Adoption	-0.5	0.4	-0.24	0.18	0.24	-0.75	-0.17
3 Years Post-Adoption	-1.27	0.81	0.17	-0.17	0.21	-1.01	-0.08
4 Years Post-Adoption	-1.58	0.8	0.67	-0.92	-0.41	-1.21	-0.11
5 Years Post-Adoption	-1.92*	1.74*	0.74	-1.24	-0.3	-1.3	0.14
Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**-2.34****	**2.44****	**1.59**	**-2.61*****	**-0.81**	**-1.57**	**-0.01**
Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**-0.87**	**-1.72***	**-1.55**	**0.96**	**-0.56**	**1.38**	**0.23**
Variance 2 Years Pre-Adoption to 5 Years Post-Adoption	1.24	-0.14	-0.48	-1.42	-1.15	-0.32	0.04
Variance 7 Years Pre-Adoption to 2 Years Pre-Adoption	1.32	0.29	-0.16	0.38	1.04	-3.39***	-0.17
Treated Observations	21	32	11	33	32	49	61
Untreated Observations	118	238	75	248	241	419	553

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

***Notes*: T statistics above for mean differences, treated (adopters) minus untreated (matched controls), all measures are % of GDP. Propensity scores for matching are generated by hazard regression on Table II-2.**

Table II-5
Comparison of Means without Matching

	(1) WEO Central Government Debt	(2) WEO Surplus/Deficit	(3) WEO Primary Surplus/Deficit	(4) WEO Expenditure	(5) WEO Tax Revenue	(6) WDI External Debt	(7) MF External Debt
7 Years Pre-Adoption	3.26***	-0.26	-1.04	-3.23***	-3.65***	0.59	-0.03
6 Years Pre-Adoption	3.48***	-0.33	-1.44	-3.09***	-3.51***	0.14	-0.31
5 Years Pre-Adoption	3.29***	-0.63	-1.01	-2.94***	-3.59***	0.16	-0.4
4 Years Pre-Adoption	3.25***	-0.6	-1.35	-3.01***	-3.59***	-0.03	-0.63
3 Years Pre-Adoption	3.32***	-0.65	-1.21	-2.73***	-3.2***	-0.21	-0.76
2 Years Pre-Adoption	3.45***	-1.87*	-2.82***	-1.83*	-3.12***	-0.39	-0.95
1 Year Pre-Adoption	2.82***	-1.3	-2.42**	-1.85*	-2.76***	-0.64	-1.04
Year of Adoption	1.58	-1.76*	-2.15	-1.84*	-3.02***	-1.26	-1.2
1 Year Post-Adoption	1.93*	-1.62	-1.69*	-2.25**	-3.11***	-1.31	-1.17
2 Years Post-Adoption	1.87*	-1.28	-1.36	-2.42***	-2.84***	-1.47	-1.35
3 Years Post-Adoption	1.28	-1.14	-1.47	-2.53***	-2.67***	-1.61	-1.22
4 Years Post-Adoption	1.18	-0.94	-1.41	-3.05***	-2.93***	-1.85**	-1.2
5 Years Post-Adoption	0.86	-0.39	-1.23	-3.07***	-2.73***	-2.01**	-0.93
Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**-3.55*** **	**1.81***	**2.06****	**-0.9**	**0.6**	**-1.58**	**-0.06**
Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**-1.86***	**-1.76***	**-2.43*****	**1.99****	**0.38**	**-1.29**	**-0.92**
Variance 2 Years Pre-Adoption to 5 Years Post-Adoption	2.83***	-2.17**	-2.05**	-2.64***	-2.5***	-0.93	-0.83
Variance 7 Years Pre-Adoption to 2 Years Pre-Adoption	3.5***	-1.69*	-1.31	-1.58	-1.4	-1.38	-0.85
Treated Observations	21	32	11	33	32	49	61
Untreated Observations	118	238	75	248	241	419	553

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

***Notes*: T statistics above for mean differences, treated (adopters) minus untreated (non-adopters), all measures are % of GDP. Propensity scores for matching are generated by hazard regression on Table II-2.**

Table II-6
Comparison of Matching Methods, Levels

Variable, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	Single Nearest Neighbor, No Replacement	Three Nearest Neighbors	Single Nearest Neighbor	Caliper .1	Caliper .05	Epanechnikov Kernel	Total Significant 10% level
WEO Central Government Debt	-3.22***	-2.70***	-1.84*	-2.31**	-2.13**	-2.13**	7
WEO Surplus/Deficit	2.77***	3.64***	2.37**	3.46***	3.72***	3.69***	7
WEO Primary Surplus/Deficit	2.83***	2.31**	2.67**	1.49	3.16***	3.16***	5
WEO Expenditure	-3.23***	-3.06***	-3.54***	-2.31**	-2.85***	-2.79***	7
WEO Tax Revenue	-0.89	-0.59	-1.67	0.12	-0.41	-0.41	0
WDI External Debt	-0.82	-1.87*	-0.67	-1.82*	-1.56	-1.77*	3
MF External Debt	0.77	0.37	-0.08	0.18	0.47	0.13	0

Variable, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	Single Nearest Neighbor, No Replacement	Three Nearest Neighbors	Single Nearest Neighbor	Caliper .1	Caliper .05	Epanechnikov Kernel	Total Significant 10% level
WEO Central Government Debt	-1.15	-1.26	-0.67	-1.21	-1.19	-1.17	0
WEO Surplus/Deficit	-2.90***	-2.39**	-2.19**	-2.12**	-2.31**	-2.33**	7
WEO Primary Surplus/Deficit	-1.48	-2.15**	-1.55	-2.29**	-2.59**	-2.58**	4
WEO Expenditure	1.67	1.95*	1.52	1.1	1.55	0.02	1
WEO Tax Revenue	-0.43	0.02	-0.23	-0.39	-0.23	-2.37**	1
WDI External Debt	1.24	-0.15	-0.03	1.82*	0.36	0.10	1
MF External Debt	0.31	-0.38	-0.15	0.51	0.01	-0.00	0

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

***Notes*: T statistics above for mean differences, treated (adopters) minus untreated (non-adopters).**
Propensity scores for matching are generated by hazard regression on Table II-2. Total significant also counts results from the normal kernel on Table II-4.

Table II-7
Comparison of Matching Methods, Variance

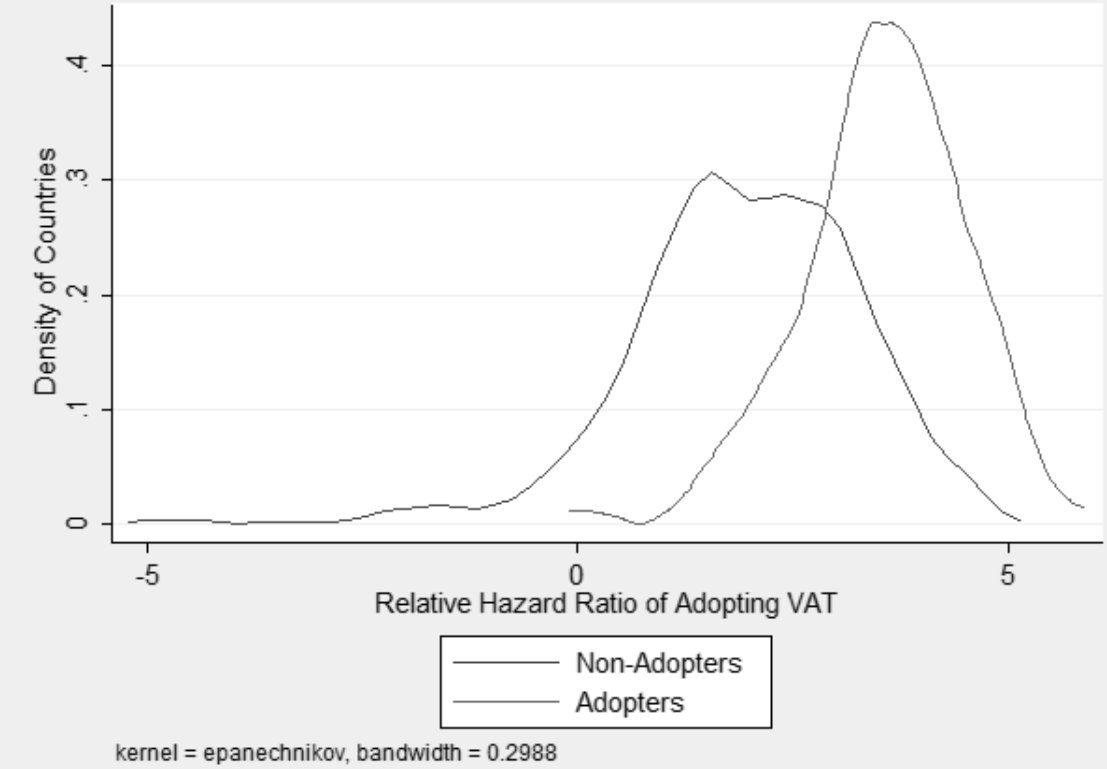
Variable, 2 Years Pre-Adoption to 5 Years Post-Adoption	Single Nearest Neighbor, No Replacement	Three Nearest Neighbors	Single Nearest Neighbor	Caliper .1	Caliper .05	Epanechnikov Kernel	Total Significant 10% level
WEO Central Government Debt	1.19	1.21	1.17	1.25	1.03	1.03	0
WEO Surplus/Deficit	0.72	-0.11	0.41	-0.42	0.11	0.15	0
WEO Primary Surplus/Deficit	-1.48	-0.75	-1.66	-1.28	-1.47	-1.49	0
WEO Expenditure	-2.38**	-2.46**	-2.42**	-1.94*	-2.11**	-1.90*	6
WEO Tax Revenue	-1.45	-1.11	-1.73*	-0.40	-0.73	-0.70	1
WDI External Debt	-1.28	-0.69	-0.60	-0.33	0.30	-0.21	0
MF External Debt	0.80	0.82	0.84	0.42	0.28	0.31	0

Variable, 7 Years Pre-Adoption to 2 Years Pre-Adoption	Single Nearest Neighbor, No Replacement	Three Nearest Neighbors	Single Nearest Neighbor	Caliper .1	Caliper .05	Epanechnikov Kernel	Total Significant 10% level
WEO Central Government Debt	1.41	1.30	1.35	1.40	1.41	1.42	0
WEO Surplus/Deficit	0.88	0.22	0.90	0.46	0.50	0.55	0
WEO Primary Surplus/Deficit	-0.01	0.08	-0.42	-0.25	-0.29	-0.29	0
WEO Expenditure	0.99	0.51	0.77	0.68	0.32	0.31	0
WEO Tax Revenue	1.51	1.32	1.31	1.35	1.09	1.12	0
WDI External Debt	-2.59***	-1.32	-0.75	-5.16***	-2.80***	-2.28**	5
MF External Debt	-1.81*	-0.69	-0.48	-0.57	-0.59	-0.43	1

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$

***Notes*: T statistics above for mean differences, treated (adopters) minus untreated (non-adopters).**
Propensity scores for matching are generated by hazard regression on Table II-2. Total significant also counts results from the normal kernel on Table II-4.

Figure II-1
Evidence of Common Support- Kernel Density Plot



III. The Impact of English Language Skills on National Income: A Cross-National Comparison

1. Introduction

While education and literacy have received a great deal of attention in macroeconomics, languages receive far less attention in the role they play in economic development (Arcand and Grin 2012.) This subject deserves far more study, as a common language between two countries may raise incomes in each by facilitating trade and foreign direct investment via lower transaction costs as well as improve access to cost reducing technology. In addition, better foreign language abilities often open up more lucrative employment opportunities both within the country and for those emigrating (Lee 2012, Angrist and Lavy 1997, Lein et al 2011, Melitz 2007, Meltiz and Toubal 2014.) Language specific transaction costs are more important for services than for manufacturing, as services generally imply more interpersonal interactions. This suggests that the structural changes inherent in the development process, moving from agrarian to manufacturing to service industries as countries grow richer, may reinforce the importance of improving foreign language skills in the future (Warschauer 2000.)

English has generally become the de facto international language for business, advertising, academia, media, and numerous other fields (Nunan 2003, Crystal 2003.) As a result, many developing countries, recently including Rwanda and Madagascar, have begun to integrate English as a part of school curriculums, either as standalone subjects or by changing the language of instruction to English with the goal of creating a better prepared workforce for international industries. At a micro level, citizens often work to improve their skills in English in hopes of raising their wages, such as seeking better employment in the service sector like call centers (Nunan 2003, Casale and Posel

2011, Chakboraty and Kapur 2009, Birdsall 2011.) USAid, the US governmental agency responsible for foreign development work, promotes English education in a number of developing countries. In a world with ever increasing international competitive pressures, English abilities are proposed to be an important component of participating in the global economy (Wedell 2008, and Warschauer 2000.)

However, the benefits from better English skills are currently not well understood. While micro studies find individuals with better language skills earn more, macro evidence is scarce. Higher English abilities could improve outcomes operate through two broad channels. The first channel, through trade, eliminates the need for translation or language training that would make it costly to export internationally or attract investment from foreign multinationals. The second, through technology, gives people access to better information, allow them to use lower per unit cost technologies. While benefits of learning English to an individual are readily identified, usually higher paying employment, the costs may prove to be too high for a social planner considering education in language skills versus other forms of human capital (Wedell 2008.) The focus on English language skills may be to the detriment of other useful skills, and this may negatively affect poorer students with the fewest opportunities to put their English skills to use (Bruthiaux 2002.) While having a shared language definitely increases trade volume, which hopefully leads to higher incomes, it is not clear that English itself plays a special role in international trade. Hejazi and Ma (2011)finds that two countries that both speak English have more bilateral trade than two countries who share a different language, while Melitz (2007) and Melitz and Toubal (2014) find that English plays no special role compared to other languages.

Empirically identifying the impact of language skills is problematic, as it is not clear if English language skills lead to higher incomes and trade, or higher incomes and trade leads people to study English. From a policy perspective, English education is costly and it should be promoted if it can be shown to most cost effectively improve outcomes, as some micro level studies have shown; otherwise, resources may be better spent elsewhere. To answer whether on a macro level English skills improve incomes and employment opportunities, I employ an index of English language skills, the Test of English as a Foreign Language, or TOEFL, to estimate the impact of English language proficiency on income per capita, net exports, FDI and emigration over the period of 1992-2012 for 139 countries.[23] I use an instrumental variables framework to obtain a consistent estimate of English skill's effects. The instrument, a measure of linguistic distance, is hours required to learn English conditional on a country's official language. This measure is correlated with English skills as it represents how much effort would be required to learn English. However, aside from this channel, linguistic distance is not correlated with any of the outcome variables of income, net exports, FDI and emigration.

I find that while the impact of English abilities on income and net exports is positive using OLS, the estimated effect is even larger after using instrumental variables. This relationship is robust to several different ways of parameterizing the instrumental variables. There is no such observed effect of language abilities on emigration or FDI. This indicates that English abilities are correlated with higher

[23] Data availability varies by year: for the 2012 year that gets a large portion of the analysis, 139 countries have data available on TOEFL scores. 120 are available in 1992, the start of the period. 173 countries have TOEFL data available for some years. Data is available from the author upon request.

incomes in part through higher net exports, and this is not driven by foreign investment. This would support the story that domestic industries are better able to find profitable export markets for their products and services when their employees have higher English abilities, or countries with higher English abilities are able to go into more profitable export-based industries. Better English abilities may also allow countries access to lower cost technology. This shows the promise of English language education in a positive light, as higher English abilities opens countries to more markets and ultimately higher standards of living.

This paper is organized as follows: section I outlines the goals of this paper, section II introduces the relevant literature, section III discusses the estimation strategy, section IV discusses the data, section V discusses results, and section VI concludes.

2. Literature Review

2A. English: Benefits and Costs

USAid in particular promotes English education in its development work, in recent years advertising its English education programs as success stories in development. For example, engineers in Kabul taught English are now able to employ electronic control equipment, whose use requires one to know English, instead of manual control equipment to run the local hydroelectric power plant. With their better English abilities, they are also able to seek specialized education abroad. Because of access to better education and technology through higher English skills, there are now fewer accidents and service outages (USAid 2013.) This argument would lend credence to the idea that higher English abilities would allow workers in the developing world to

become more productive, as they are able to use better equipment and have better access to information. USAid assisted Rwanda's changing of its official language of instruction from French to English, a change intended to raise incomes and improve international trade (USAid 2014.) USAid claims its English training programs in countries such as Ethiopia and the Philippines are a means of improving the lives of people in these countries by broadening job opportunities, going so far as to draw fire from US lawmakers fearing the development program will lead to some service jobs such as call centers to locate in these developing countries instead of the US (Koch 2014, De Lotbiniere 2012, May 2012.) English education is thus considered by some development agencies as a useful and effective tool, and powerful enough to arouse the concern of protectionist lawmakers.

Despite all these promised benefits, the true effects of higher English language abilities and English language education are unclear. There is mounting evidence that countries promoting instruction in English may also be exacerbating income inequalities. Glewwe et al. (2009) found that the use of English-language textbooks in primary school was of limited use to most students in Kenya. Students lacked the ability to read English language textbooks and only high achievers and wealthier students saw any benefit from receiving these textbooks. Bruthiaux (2002) warns that a focus on English language education may worsen income gaps in the developing world. Only the rich will have the resources to benefit from learning English or being instructed in English, as the poor would never have access to the networks and international opportunities that would make learning English worthwhile. Poorer students may have better earning potential if they received instruction in their native language, with more

emphasis on non-language related subjects such as mathematics, science, trades, or writing skills in their native language. This highlights a tradeoff between more general human capital, and English specific human capital, as well as the conflict between promoting a more internationally useful language against one that would preserve local culture and prepare students for domestic-focused employment (Prah 2008.)[24]

2B. Effects of Language: Micro

There is some literature on the microeconomic returns to language skills, specifically English, on income, that will guide the empirics of this paper. Their findings and methodology are difficult to generalize from an individual to a national level due to data availability, data heterogeneity, and choice of control variables. There also exist several papers using gravity models to study the impact of common languages on trade that, while taking a different approach to a similar question, are instructive with their choice of control variables and parameterization of linguistic similarity.

Several microeconomic studies have found a positive impact of English skills on individual earnings in India and South Africa (Azam et al 2013, Casale and Posel 2011.)

[24] Many developing nations face the challenge between teaching their native languages, or teaching a more widely spoken foreign language, often English (Bambgbose 2009, Wolhuter et al 2006.) Wedell find that educators worldwide are dissatisfied with this tradeoff, being forced to choose between subjects without adequate information about their payoffs (2008.) In order to improve literacy rates in sub-Saharan Africa, some countries have found success in switching from teaching a colonial language- French or English- instead instructing in in the native language of students (Bamgbose 2009.) English language skills deserve more empirical study as educators and policymakers alike want to maximize opportunities of students through the choice of curricula. Due to these expensive tradeoffs and even potential downsides to English instruction and study of the English language, the recent moves of some countries, such as the Gambia, away from English are now far more understandable.

Other studies have considered immigrants in the United States (Kossoudji 1988.) However, these studies and likely any study comparing English skills and incomes are inherently complicated by endogeneity problems, as language abilities affect income and income affects language abilities. Lang and Sinvier show that returns to English language skills in the form of higher incomes, even in countries where English is not an official language (2006.)

This pattern is not limited to the English language. Languages may affect trade or employment opportunities, as seen in Angrist and Lavy's (1997) study of French in Morocco and Rendon's (2007) study of Catalan in Catalonia. Angrist and Lavy's paper use an instrumental variables regression to explore the switch from French to Arabic instruction as the language of official instruction in Morocco. This change in instruction reduced student's French writing skills, which in turn limited student employment opportunities and as a result lowered income. Silvio created a theoretical model of language learning, which demonstrated that knowledge of more languages led to higher paying job opportunities in Catalonia. These studies are microeconomic in nature, and so a macroeconomic study may find different results even if one uses a similar methodology. For example, microeconomic literature going back at least as far as Mincer (1974) generally find positive returns to education on an individual level, but macro studies often finds no effect of countries increasing average levels of education on incomes (Pritchett 2001.)

2C. Effects of Language: Trade

Other papers have considered the interaction between language skills and bilateral trade. Egger and Lassman (2013) use a case study in Switzerland to find that areas with more linguistic similarity tend to trade more, ascribing this effect to cultural similarity. Lein et al. (2011), using a gravity model, find that shared languages lead to both an increase in bilateral trade and FDI inflows. They propose choosing to study a language is an optimization decision where participants maximize benefits of such education conditional on its costs. Molnar (2013) looks at the role translation costs play in international trade, finding a significant effect on bilateral trade. Finally, Egger and Lassman's (2012) survey and meta-analysis of the literature consistently finds that linguistic similarity is a major determinant of trade, even once controlling for factors such as legal origin, colonial history, and exchange rates.

Looking more directly at the role played by English, Hejazi and Ma (2011) employ a gravity model to find if English has a premium as a shared language over other common languages among OECD countries. They find results that agree with previous authors in that shared languages are positive in their effect on trade, but that sharing English has an even greater effect. Melitz (2007) divides linguistic similarity into an extensive, whether the country pair has a shared language, as well as intensive margin, how many people in those countries share that language. While both effects are positive, the intensive margin's effect is stronger. However, while shared languages matter, and shared European languages as a whole have a stronger effect on bilateral trade, English itself has no larger positive marginal effect on bilateral trade than other European languages. Melitz and Toubal (2014) extend this paper, using a more

complex measure of linguistic proximity, and find similar results. Thus, while common languages raise bilateral trade, whether English plays a special role is unclear.

2D. Further Considerations

So far, such a comprehensive study of English language skills and education worldwide has been hampered by lack of data. While an international database on English language education would be ideal, with information on time and money committed to the subject of English language study by country as well as their languages of instruction, it currently does not exist. A few regional surveys exist, such as the one Nunan (2003) performed in East Asia. Even if the data did exist, such cross-national comparisons are likely limited due to heterogeneity in education as well as data quality. Studies looking at the impact of English education may thus be limited to considering only one country.

Due to these limitations, this paper considers national English language skills rather than English language education. This allows me to evaluate the value of the end product of English language education, the stock of English language skills.[25] To this end, I use data from the TOEFL, a measure of academic English language skills, to proxy for national level English language abilities. I will discuss estimation of these effects, the challenges posed, and possible solutions in the next section.

[25] Whether a particular education program or method produces the most human capital per dollar spent is beyond the scope of this paper, as this paper only considers the average level of English language based human capital present in a country.

3. Estimation Strategy

In order to obtain consistent estimates of the impact of English skills on macroeconomic outcomes, one must solve the problem of endogeneity: Y_i, most importantly per capita income but also net exports, emigration, and FDI in this paper, is a function of English ability E_i. This higher ability enables easier communication and translation that could raise trade, better access to technology that raises incomes, make a country more attractive to foreign investors attempting to open a call center or other industry, or lead to more emigration of people seeking to use their language abilities abroad. However, English abilities are also a function of income as well as other outcome variables. Countries with higher income likely have an easier time improving their language abilities, and inhabitants of countries already experiencing a great deal of trade or foreign investment are likely to seek to improve their language abilities. Alternatively, very poor countries may be the only ones selecting to improve English abilities.[26] Both the outcome variables and language abilities are likely influenced by some historical elements, and the process is likely to be self-reinforcing, with changes in one also affecting the other. The model can be described as follows:

$$(1) \qquad Y_i = A_0 + A_1 E_i + A_2 X_i$$

$$(2) \qquad E_i = B_0 + B_1 Y_i + B_2 X_i$$

Where X_i is a vector of common exogenous covariates. Using OLS to estimate for income per capita as Y_i, the estimate A_1will be upwardly biased and inconsistent

[26] Because the cost of learning a language is in large part time, the opportunity cost for people to learn a language in poor countries with poor employment opportunities may actually be lower than that of wealthier countries relative to their outside options. People have after all become multi-lingual throughout history, when even wealthier people by that era were quite poor by modern standards.

due to this endogeneity if there is self-reinforcement with richer countries focusing more on English abilities, or downwardly biased if improving English abilities is a treatment only undertaken by already poorer countries. Trying to estimate A_1 consistently will thus require a different estimation method, such as instrumental variables.

3A. Instruments

Previous research on the impact of language skills on individual level income has employed instrumental variable estimation. Akbulut-Yuksel et al (2011) as well as Bleakley and Chin (2004) employ an individuals' age of immigration as an instrument for English ability, as younger children are better able to learn a new language than adults, and age of immigration itself should have no direct impact on income. Dustmann and Van Soest (2002) as well as Rooth (2001) employ parental education as an instrument, as it should be correlated with language ability but not directly with income. Chiswick and Miller (1995) employ country of origin and marriage status as instruments, as married immigrants with a spouse who speaks the same language will have less incentive to learn the language of his new country. However, marriage status and country of origin may still be correlated with income. Finally, Chakboraty and Kapur (2009) use a policy change on English instruction in an Indian state as an instrument, and such an experiment does not exist for all countries. As this paper considers cross national variation rather than cross individual variation, different instruments will be needed from those previously considered, as these instruments cannot be generalized to a national level.

A more feasible class of instrument for this paper, suggested by Shastry (2007) and also implemented by Isphording and Otten (2014), is linguistic distance. Shastry considers the fact that many provinces in India speak languages very linguistically different from Hindi, and others speak ones that are more similar. Learning Hindi is more common in the latter group, as the costs are lower, while learning English is more common in the former, as the benefits are higher for the same level of effort. Worldwide, some languages are linguistically close to English and this makes it easier for native speakers to learn English compared to native speakers of other languages. Thus, one's native language should raise the time required to learn English and this increase in the time required to learn English should be exogenous. As a result, linguistic distance from English should be negatively correlated with English language skills.[27] Isphording and Otten (2014) find that this increased difficulty of learning English can be considerable depending on their original languages and leads to easily observable differences in incomes of immigrants. However, these studies can be hampered by a lack of concrete measure of distance, with Shastry (2007) relying on estimates of how technically different languages are to learn as classified by linguists

[27] There is no reason to expect linguistic distance from English will impact income or trade other than through the channel of ease of learning English, as there is nothing inherently different between languages' ability to communicate ideas or structure that should affect income. Crystal (2003) argues that the supremacy of English in international trade, business and media is entirely accidental due to the course of history. It is plausible that an alphabet-based language, like English, is superior to a pictograph based one, such as Chinese, due to the ease of which it can change and add new ideas. But Romanization of Chinese through pinyin, for example, illustrates that a pictographic language can be converted to an alphabet based one. Additionally, as children learn to speak their first language at roughly the same age, from an equal starting point no language is more difficult to learn or communicate with than any other (Crystal 2003.)

and Isphording and Otten (2014) employing a method of linguistic similarity based on comparing grammatical structure and cognates.

In this paper, I employ a source that gives definitive quantitative measures of difficulty for native speakers of one language to learn English. The Foreign Service Institute (FSI) produces information on how many hours of instruction are required for a native English speaker to learn a foreign language, and assuming this time is symmetric, as assumed by Chiswick and Miller (1999), it will also give an unbiased estimate of the time it would take a given native language speaker to learn English. Languages are classified in 7 categories of difficulty, going from requiring 600 hours of study to over 2200 hours of study, with a summary of hours provided on Table III-1. Adding in linguistic distance, D_i the estimation of the earlier system of equations (1) and (2) thus becomes:

$$(3) \quad Y_i = A_0 + A_1 E_i + A_2 X_i$$

$$(4) \quad E_i = B_0 + B_1 Y_i + B_2 X_i + B_3 D_i$$

While one cannot identify the English language equation (4), one can identify the outcome equation (3), using linguistic distance as an instrument for English language skills to obtain a consistent estimate of their impact on per capita income, trade as a share of GDP, or emigration.

This instrumental variables approach's advantage over previous approaches is that it uses a commonly employed cost- measured in hours- to quantify linguistic distance, rather than assumed difference based on essentially untestable assumptions about how similar languages are at an abstract level. At the same time, the instrumental variable clearly satisfies the exclusion restrictions. The difficulty of learning English

will definitely affect one's decision to learn the language, as it will change the relative cost of accumulating other forms of human capital or seeking other forms of employment instead of studying English. It will have no effect on the outcome variables, as the difficulty one has in learning English should not affect income, trade, foreign investment, or emigration, aside from one's actual English abilities.

However, the proposed functional form of this cost instrument requires two assumptions that may be considered restrictive. First, while the costs of learning a language may not be linear, the instrument would require this relationship to be linear, or at least require a fixed functional form. For example, in a linear form, languages that take 800 hours to learn require half the resources of a language that takes 1600 hours to learn. This is not necessarily true, as individuals may not place a linear value on time. Second, the assumption of symmetry may be inaccurate due to the complexity of English. While an English speaker may have an easy time learning Spanish in 600 hours, a Spanish speaker may take more hours to learn English. However, the difficulty of learning English would thus come from two sources- its overall difficulty as a language, which should be a constant for all people trying to learn it regardless of one's native language, and its closeness to another language, which varies based on one's native language. Spanish speakers should thus find it easier to learn English than Mandarin speakers, which should preserve the ordinal rank of the time required to learn English for other language speakers, even if it does not preserve the cardinality of difficulty.

Both of these concerns can be addressed by using a series of dummy variables for the different categories of language difficulty instead of the hours required to learn

them, as this does not restrict the linguistic distance to be linear in hours required to learn English, or require it to take the same number of hours to go from English to Spanish as it does to go from Spanish to English. I parameterize this instrumental variable in two ways.. In one specification, each of the seven categories on Table III-1 receives their own dummy variable, with native English speaking countries being the eighth category omitted for perfect co-linearity. This set of dummies is then used as the instrument. However, many of these categories have very few languages or include very few countries, so I also employ specification that reduces the number of categories dummy variables to three (four including omitted native English for perfect co-linearity.)[28] As a robustness check, I use a specification created from a separate measure of linguistic similarity based on technical proximity of words, the Linguistic Proximity-2, or LP2 variable, used by Melitz and Toubal (2014.) Higher values of this variable represent more linguistic proximity, and lower values less proximity.

Table III-2 presents basic correlations that are consistent with these instrumental variable exclusion restrictions for the 20 years of data available, 1992-2012 excluding 2011.[29] TOEFL is positively correlated with natural log per capita real GDP at .3968, and TOEFL scores are negatively correlated with hours to learn a language at -.2837. Hours to learn English are negatively correlated with per capita income at .0376, extremely close to zero. This evidence would support omitting linguistic distance from the income equation, and suggesting it is an appropriate instrument for English abilities

[28] The first dummy category is equal to 1 if a native language speaker of a country takes 600, 750, or 900 hours (first three categories of Table III-1) to learn English and 0 otherwise, the second 1 if it takes 1100 or 1100+ (fourth and fifth categories) to learn English and 0 otherwise, and the third category 1 if it takes 2200 or 2200+ (sixth and seventh categories) to learn English.

[29] Data is not available for 2011.

in that same equation. LP2 has a stronger correlation with score than hours to learn, .4972, but also is highly correlated with the outcome of income at .5382. This would indicate that this alternative instrumental variable used by other authors may not satisfy exclusion restrictions in this case.

4. Data

This paper expresses English abilities through the national average performance on the TOEFL exam. The TOEFL is an assessment of a student's ability to use English in the classroom, considering four skills: reading, listening, speaking and writing. 27 million students having taken the test to date in over 4500 testing sites and across 165 countries since it was organized in 1964, giving a broad base of support for the test scores. TOEFL exams were given as a paper test for much of its history, switching to local computer based examinations in mid-1998 and then an internet based examination in mid-2005.

The most recent version of the TOEFL has a number of advantages. A major strength of the TOEFL is that it allows the scores to be comparable internationally. For internet based tests, students should be facing roughly the same testing environment; although paper or local computer based testing students may face a different environment. It has a broad history, with data available from 1992-2012 for most countries, allowing one to construct a relatively balanced panel or observe variation over time. Finally, the examination is generally accepted to be the best assessment of English language abilities that is widely available, as measured by the number of institutions who accept it. TOEFL scores are accepted by over 9000 colleges. This is

more than its largest alternative, the International English Language Test IELTS at 8000, which also has a shorter history, fewer test takers, fewer testing centers, and presence in fewer countries, although the difference is relatively small.[30] TOEFL scores are the best proxy of general English abilities available for a country.

The TOEFL index has several shortcomings, none of which are critical. First, there is the incentive to cheat or otherwise compromise the integrity of the test. TOEFL scores are usually a major requirement for admission to an English speaking university and so high scores in a country may either represent higher ability or better cheating. This would introduce an attenuation bias, biasing towards the effects of English language abilities on income or other outcomes towards zero. If one observes a positive effect, it is possible the true effect is even higher as the current on is observed with attenuation from cheating. However, cheating should be more difficult on internet based tests than on paper exams, where the central testing body has more oversight on how the test is executed and how its results are transported for evaluation.

Second, while test takers may self-select, it is not clear how selection would bias results in this setting. It is possible only more competent English speakers will be motivated to take an examination, or that only less competent speakers will take the test while more competent ones have advanced positions that do not require testing. However, selection to take the exam is likely determined on an individual rather than national level, and this study employs national aggregates. Thus, even if only the best English speakers in each nation select to take the exam, one would still be comparing the scores of the best individuals across nations, which would also be the individuals

[30] IELTS also does not release data publically

most relevant to conducting international business interactions or making use of cutting edge technology. A variety of skill levels are present in the index, and most countries are represented, so it appears countries with poor language abilities are not avoiding the exam altogether.

Finally, the index shows the average English proficiency of a test taker, a quality measure, not the share of inhabitants that speak English, a quantity measure. However, if one can determine that average English proficiency is correlated with share of inhabitants that speak English, one could claim that the TOEFL is a good summary of both quantity and quality of English speakers in a country. Data on English speakers are available from a number of sources such as national censuses as well as Crystal (2003). Using this data, I can compare English speakers as a share of population to English proficiency for about half the countries with TOEFL data available. Table III-2 shows correlations that indicate it is indeed the case that the share of population that speaks English is correlated with the TOEFL at .3321.

Table III-3 presents summary statistics for the variables used in the regression. Note that there is a reasonably large range of abilities for the TOEFL. Countries that have multiple official languages with different linguistic distances from English according to the FSI use the value of the language most distant from English among all their official languages for this paper. Additionally, following Meltiz and Toubal (2014), I include a control dummy, equal to 1 for a country having English as one of their official languages, in order to capture the effect of official endorsement or presence of English aside from the actual abilities of a country.

Left hand side variables include net exports (exports minus imports as a share of GDP). FDI(net inflows from foreign countries as a share of GDP), natural log of real GDP per capita, and emigration rate data from the OECD from 2000-2008 across different specifications. This emigration data includes 200 countries of origin and 38 OECD and affiliated countries as destinations, but is not available as a panel.

For covariates, a natural and important inclusion would be an overall measure of education. There is a great deal of interaction between language ability and overall education level, with well-educated individuals best being able to leverage the benefits of English in other studies (Azam et al. 2013.) Barro and Lee (2013) produce a measure of overall education, updated in five year windows, which I employ for this study. I use a weighted average of this education measure for the years where measures of education are not present, so if education values are available 1995 and 2000, 1996 uses 20% of the value from 2000 and 80% from 1995, 1997 uses 40% of the value from 2000 and 60% from 1995, and so on.

Persistently high inflation may negatively affect real incomes, so I include inflation as an independent variable. Government size as a share of GDP may also impact long term growth, so this is also added as a right hand variable. As I am analyzing linguistic factors, it is sensible to include a measure of ethnolinguistic fractionalization, as countries with more languages may have lower incomes as well as face different incentives to learn English (Arcand and Grin 2012.)

Geographic factors such as latitude, landlocked status, or island status may also affect incomes or the other outcomes being considered. Easterly (2001) and Bloom et al (1998), among others, consider these geographic features to be potential determinants of

income. The size of the country may also matter, as larger countries have lower incentives to learn non-native languages as they have many more commercial opportunities within their borders, so I include natural log of population as well as land area as a covariate. I also include continent level dummies with Europe being the omitted continent due to perfect co-linearity.

Ethnolinguistic fractionalization comes from Roeder (2001), latitude data is from the CIA World Factbook, and continents are specified by the ETS in their TOEFL publications. All other covariates come from the World Development Indicators. The linguistic distance instrument, hours to learn English as listed by the FSI, is parameterized in two different ways as discussed earlier. I also employ Meltiz and Toubal's (2014) LP2 measure as a robustness check.[31]

5. Results

5A. Income

Table III-4 shows the impact of English language abilities and various covariates on GDP per capita. The results here are given with standardized betas, where a 1 standard deviation change in the X variable leads to a B (listed in table) standard deviation change in the Y variable. The OLS specification (1) and (2) shows a positive and significant effect of English abilities on incomes, but the magnitude drops substantially once covariates are included. The only covariate value that is particularly

[31] Several countries lack official languages that are featured on the FSI list of language learning difficulty. The number of hours to learn English by nearest linguistic neighbor is used instead. Additionally, for countries with multiple official languages of varying linguistic distance from English, the language with the largest distance from English is used.

surprising is the dummy for English being an official language, which is negative. This would imply that once I control for English abilities, officially recognizing English is negative in its effect on income.

All three IV specifications, specification 1 breaking linguistic distance into three categories, specification 2 breaking it more finely into seven categories, and the LP2 measure using a continuous measure of linguistic distance, show a positive and significant effect of English language abilities on incomes after accounting for covariates. This effect is also much larger than the effect observed using OLS. Depending on the specification a one standard deviation increase in TOEFL scores raised log income by between .358 and .768 standard deviations. Presented using non-standardized coefficients, raising the average TOEFL score of a country by 1 point on a 120 point scale raises its real per capita incomes by between 6 and 13 percent.

These findings could indicate either that the IVs corrected for an attenuation effect, as TOEFL scores may not accurately capture English abilities, or that lower income countries were the ones seeing the largest benefit from better English abilities. Both stories are consistent with the increased size of the coefficient on the TOEFL score. The argument that higher income countries choose to improve, or would find it easier to improve, their English abilities is either inconsistent with this evidence or dominated by the other two effects. Otherwise, IV results should have a smaller coefficient than OLS results. The Hausman test strongly rejects the null of no difference between OLS and IV in all cases, indicating a statistically significant systematic difference between IV and OLS estimations in all IV specifications.

First stage results for IVs in Table III-4a are unsurprising given these findings. The coefficients for dummies in specifications 3-4 are steadily decreasing as the hours needed to learn English increases, indicating more linguistic distance from English leads to lower TOEFL scores. They are decreasing less steadily for specifications 5-6, possibly because several of the categories contain a small number of countries. Similarly, for 7-8, where higher values of LP2 mean higher levels of linguistic proximity, the correlation is positive.

5B. Trade

Having established that English is correlated with higher incomes, I now explore channels by which this effect may work. As argued earlier, English abilities may open more export markets due to ease of communication, and they may also give access to more productive technology. While the latter is difficult to test due to the general inability of economists to identify technology as anything beyond a residual, looking for an increase in exports specifically is possible. As higher English abilities is supposed to open more export markets due to ease of communication, one may expect higher net exports to be a natural result of higher English language abilities. Table III-5 shows exactly this, following a similar pattern as seen in Table III-4 with income. There is a positive effect of TOEFL scores on net exports both with and without covariates in the OLS specifications (1-2). All IV specifications with covariates (4, 6, 8) show the same positive, statistically significant, and statistically different from OLS (as by the Hausman test) effect of TOEFL scores on net exports as a share of GDP. This effect is comparatively large, with a 1 standard deviation change in TOEFL scores raising net

exports as a share of GDP by between .6 and 1.3 standard deviations. In unstandardized terms, a 1 point increase in average TOEFL scores on a 120 point scale raises net exports between .8 to 1.9 percentage points as a share of GDP. First stage results in Table III-5a are similar to Table III-4a. This would agree with Hejazi and Ma (2011), finding that English is significant in generating trade specifically, as well as numerous authors finding shared languages raise trade.

5C. Foreign investment and Emigration

While the evidence presented so far would promote the idea that higher English abilities are associated with higher incomes, through higher net exports from or through access to better technology directly raising productivity or incomes, other channels could be at work. International corporations may take advantage of higher English skills in these countries, leading to a boost in foreign direct investment (Lein et al 2011) or increased migration to opportunities abroad that leads to higher remittances albeit with loss of human capital from the original country (Gupta et al 2008.)

Table III-6 looks at the impact of TOEFL scores on FDI inflows as a share of GDP. For all three instrumental variable specifications as well as OLS, there is generally no effect of TOEFL scores on FDI after accounting for covariates. For specification (4), there is a negative impact of TOEFL scores on FDI, but the Hausman test fails to reject the null of no systematic difference between the estimations. There is no evidence that higher TOEFL scores leading to increased FDI. It is possible that higher incomes and net exports may come from domestic industries taking advantage of more international export opportunities as well as having easier access to more

information and better technologies, rather than from foreign investment. Foreign corporations may have criteria other than English ability in mind for locating a production facility. Table III-6a shows similar results to previous first stage estimation results.

Finally, Tables III-7 and III-8 consider the impact of English ability on emigration. Unlike previous variables, panel information for emigration is not available, so I use cross sectional aggregate migration from 2000-2008 from the World Bank. I consider two cross sections using this migration data, one with covariates at the start of the dataset, 1992, and the other using covariates at the end, 2012. Table III-7 shows the 2012 specification results which show, once accounting for covariates, there is generally no effect of TOEFL scores on emigration save specification (4), where it is negative. However, there is no systematic difference between IV and OLS estimates according to the Hausman test even in this case. Table III-7 presents similar results using a 1992 cross section of determinants instead of 2012 cross section, across the board finding no effect of TOEFL scores on the emigration stock. Tables III-7a and III-8a show first stage results which, while less significant than previous tables due to the smaller sample size, are still roughly similar in direction and magnitude as Table III-4a seen earlier.

The net result of these regressions is that English abilities are indeed statistically significantly correlated with several important outcomes. Countries with higher TOEFL scores tend to have higher incomes, with a 1 point increase in national average TOEFL scores on a 120 point scale being associated with a 6 to 13 percent increase in real per capita income. This could be due to access to better technology that raised worker

productivity, lower transaction costs, or more export opportunities raising overall incomes. There is direct evidence of the latter, with net exports as a share of GDP increasing by .8 to 1.9 percentage points for a 1 point national average TOEFL score increase. Using these parameters as baseline estimates, policy makers can decide whether the particular costs of improving English abilities are worth these stated benefits. However, when looking at other potential channels, such as FDI or emigration, there are no such effects. Higher TOEFL scores do not lead to more foreign direct investment, and they do not increase emigration. The higher income and exports likely come from domestic industries being able to better compete internationally, possibly due to lower communication costs and easier access to information and technology. This finding would also indicate against extensive brain drain from higher English abilities.

6. Conclusion

In an increasingly globalized world, a common language greatly simplifies international trade and business, making it easier to invest in foreign countries, export products, and exchange production technologies, all with the end result of raising incomes. Due to historical accident, English has arisen as this common international language. Many countries have begun to teach English to students in hopes of giving them a competitive edge, after seeing that English language skills are correlated with higher incomes and more employment opportunities. However, it is not clear that this relationship is straightforward; higher English skills may allow an individual to earn

higher income, or higher income may lead to one seeking more English skills. Any association of English language skills with income may thus be spurious.

To answer the question as to whether English language skills actually lead to higher national incomes, more exports, FDI, or emigration, this paper employs data on English test scores from an international education testing companies' records. To account for the fact that income and English proficiency are endogenous, I use the difficulty of learning English as an instrument for English language skills. The difficulty of learning English given the most common native language in a country should be correlated with prevailing English abilities in that country, but will have no effect on income after controlling for its effect on language abilities, making it an ideal instrument. There is a positive effect of English abilities on income and net exports, but not on FDI or emigration. This suggests that English abilities raise incomes and exports but not necessarily through higher foreign investment or remittances, perhaps either due to better access to technology raising productivity for all firms within the country or fundamental changes in domestic industries allowing them to find more export opportunities. However, while this paper is able to estimate some of the effects of national level English abilities, it is unable to identify the exact costs and benefits from investing in English language education. It merely provides a strong starting point to quantifying the benefits of English language abilities on average national incomes and suggesting broad channels of causation, and I leave the analysis of marginal tradeoffs and more specific channels to future papers.

Table III-1
Hours Required for an English Language Speaker to Learn another Language

Hours	600	750	900		1100			1100+	2200	2200+
Languages	Afrikaans	German	Indonesian	Albanian	Hebrew	Maninka	Slovenian	Estonian	Arabic	Japanese
	Catalan	Luxembourgish	Kirundi	Amharic	Herero	Marshallese	Somali	Finnish	Berber	
	Danish		Malaysian	Armenian	Hindi	Montenegrin	Soninke	Georgian	Cantonese	
	Dutch		Swahili	Azerbaijani	Hiru motu	Moore	Swait	Hungarian	Dzongkha	
	Faroese			Bambara	Icelandic	Ndebele	Tagalog	Mongolian	Mandarin	
	French			Belarussian	Igbo	Nepali	Tajik	Thai	Korean	
	Italian			Bemba	Jola	Nyanja	Tamil	Vietnamese		
	Norwegian			Bengali	Kaonde	Oshiwambo	Tetum			
	Papiamento			Bislama	Kazakh	Palauan	Tigrinya			
	Portuguese			Bosnian	Khmer	Pashto	Tok Pisin			
	Romanian			Bulgarian	Kikongo	Persian	Tonga			
	Spanish			Burmese	Kinyarwanda	Polish	Tongan			
	Swedish			Carolinian	Kongo	Rukwangali	Tshiluba			
				Chamorro	Kurdish	Russian	Turkish			
				Chewa	Kyrgyz	Samoan	Turkmen			
				Chichewa	Lao	Sango	Tuvaluan			
				Comorian	Latvian	Sena	Ukrainian			
				Croatian	Lingala	Serbian	Urdu			
				Czech	Lithuanian	Serer	Uzbek			
				Damara	Lozi	Sesotho	Voro			
				Fiji hindi	Lunda	Setswana	Wolof			
				Fijian	Luvale	Setu	Xhosa			
				Fula	Macedonian	Shona	Yoruba			
				Gilbertese	Makhuwa	Silozi	Zulu			
				Greek	Malagasy	Sinhala				
				Hausa	Mandinka	Slovak				

For specification 1, The 600, 750, and 900 categories constitute the first dummy variable, 1100 and 1100+ the second dummy variable, and 2200 and 2200+ constitute the third dummy variable. For specification 2, each has its own dummy variable. The omitted dummy in each is 1 for a country where the sole official language is English and 0 if not. Countries use the hours to learn of the official language with the highest hours to learn.

Table III-2
Correlations, 1992-2012

	Score	Hours to Learn	LP2	GDPPC, Ln
Score	1			
Hours to Learn	-0.2837	1		
LP2	0.4972	-0.1427	1	
GDPPC, Ln	0.3968	0.0376	0.5382	1

Correlation between percent speaking English and TOEFL score is .3321

Table III-3
Summary Statistics, TOEFL Scores and Covariates, 1992-2012

Variable	Obs	Mean	Std. Dev.	Min	Max	Time Invariant
Emigration, % of Population	3440	7.14	10.11	0.10	43	Y
Net Exports, % of GDP	3612	87.70	51.51	0.31	531.74	N
Real GDP Per Capita, Ln	3813	8.07	1.65	3.91	11.98	N
FDI, % of GDP	3579	4.84	12.91	-82.89	366.36	Y
TOEFL Score	3050	83.43	9.67	51.00	105.74	N
Hours to Learn English	4297	967.35	637.30	0	2200	Y
Percent English Speaking	2280	45.59	31.95	0.15	100	Y
Linguistic Proximity-2	3634	0.74	0.73	0	3.77	Y
Official English	4297	0.34	0.47	0	1.00	Y
Schooling, Average Years of	2880	7.51	2.86	0.93	13.18	N
Inflation, GDP Deflator	3840	43.48	549.47	-32.81	26762.02	N
Investment, % of GDP	3469	23.13	11.25	-2.42	227.48	N
Island	4297	0.26	0.44	0	1.00	N
Land Area, Ln	4232	10.82	3.03	0.69	16.61	N
Landlocked	4297	0.20	0.40	0.00	1	Y
Population, Ln	4265	15.06	2.35	9.12	21.02	N
Government Consumption, % of GDP	3740	12.20	8.92	0.93	68.19	N
Ethnolinguistic Fractionalization	3400	0.46	0.27	0	0.98	Y
Latitude	4237	25.47	17.01	0	72	Y
Europe	4297	0.24	0.43	0	1	Y
Pacific	4297	0.09	0.28	0	1	Y
Middle East	4297	0.09	0.28	0	1	Y
Asia	4297	0.15	0.36	0	1	Y
Africa	4297	0.23	0.42	0	1	Y
Americas	4297	0.21	0.40	0	1	Y

Table III-4
Determinants of Ln Real GDP per Capita, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.394***	0.078***	-0.235***	0.768***	0.267***	0.358***	1.082***	0.575***
	(23.03)	(4.99)	(-3.71)	(7.84)	(6.41)	(7.06)	(24.44)	(7.89)
Official English		-0.051***		-0.247***		-0.130***		-0.223***
		(-3.51)		(-7.37)		(-6.35)		(-8.43)
Schooling, Average Years of		0.470***		0.283***		0.394***		0.324***
		(25.77)		(7.92)		(16.91)		(11.21)
Inflation, GDP Deflator		-0.025**		-0.017		-0.022*		-0.019
		(-2.27)		(-1.08)		(-1.82)		(-1.41)
Investment, % of GDP		0.045***		0.105***		0.069***		0.095***
		(3.71)		(5.71)		(5.12)		(5.84)
Island		0.094***		0.093***		0.094***		0.102***
		(6.63)		(4.84)		(6.21)		(5.91)
Land Area, Ln		-0.018		0.137***		0.045*		0.088***
		(-0.95)		(4.04)		(1.93)		(2.97)
Landlocked		-0.184***		-0.195***		-0.188***		-0.184***
		(-14.40)		(-11.20)		(-13.84)		(-12.14)
Population, Ln		-0.046**		-0.161***		-0.092***		-0.121***
		(-2.40)		(-5.29)		(-4.23)		(-4.39)
Government Consumption, % of GDP		-0.133***		-0.110***		-0.124***		-0.114***
		(-11.09)		(-6.61)		(-9.60)		(-7.71)
Ethno-linguistic Fractionalization		-0.031**		-0.046**		-0.037**		-0.037**
		(-2.19)		(-2.38)		(-2.45)		(-2.04)
Latitude		0.123***		0.029		0.085***		0.051*
		(5.32)		(0.84)		(3.33)		(1.71)
Pacific		0.006		0.074***		0.033**		0.068***
		(0.40)		(3.36)		(2.06)		(3.53)
Middle East		0.069***		0.320***		0.170***		0.257***
		(4.06)		(7.67)		(6.81)		(7.63)
Asia		-0.189***		0.008		-0.109***		-0.042
		(-9.81)		(0.20)		(-4.44)		(-1.35)
Africa		-0.104***		0.107**		-0.018		0.056
		(-3.90)		(2.30)		(-0.58)		(1.41)
Americas		-0.018		0.016		-0.004		0.017
		(-0.90)		(0.57)		(-0.20)		(0.67)
N	2885	2324	2885	2324	2885	2324	2707	2241
R^2	0.155	0.713	-0.240	0.467	0.139	0.672	-0.312	0.590
F	530.6	336.2	13.73	184.3	41.05	296.5	597.0	230.8
Underid F Stat P-Value		0.000	0.000	0.000	0.000	0.000	0.000	
Hausman P value			0.000	0.000	0.000	0.006	0.000	0.000

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

Table III-4a
Determinants of Ln Real GDP per Capita, First Stage Results on TOEFL Scores, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.062	0.387***				
			(1.43)	(9.17)				
1100-2200 Hour Dummy			-0.030	0.242***				
			(-0.68)	(5.74)				
2200 Hour and Up Dummy			-0.306***	0.165***				
			(-8.53)	(3.83)				
600 Hour Dummy					0.012	0.340***		
					(0.31)	(8.45)		
750 Hour Dummy					0.254***	0.229***		
					(11.38)	(10.66)		
900 Hour Dummy					-0.039*	0.187***		
					(-1.69)	(8.06)		
1100 Hour Dummy					-0.030	0.248***		
					(-0.72)	(6.24)		
1100+ Hour Dummy					-0.010	0.122***		
					(-0.43)	(5.09)		
2200 Hour Dummy					-0.292***	0.229***		
					(-8.55)	(5.52)		
2200+ Hour Dummy					-0.115***	-0.123***		
					(-6.35)	(-6.83)		
Linguistic Proximity-2							0.497** *	0.264** *
							(29.81)	(12.68)
Covariates			No	Yes	No	Yes	No	Yes
N			2885	2324	2885	2324	2707	2241
R^2			0.107	0.509	0.172	0.541	0.247	0.523
F			115.3	125.9	85.46	117.6	888.5	143.3

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

Table III-5
Determinants of Net Exports as a share of GDP, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.106**	0.214***	-0.272***	1.289***	-0.002	0.903***	0.278**	0.602***
	(5.73)	(8.77)	(-4.52)	(8.82)	(-0.04)	(10.61)	(7.02)	(6.04)
Official English		-0.207***		-0.507***		-0.399***		-0.324***
		(-9.17)		(-10.12)		(-11.61)		(-8.96)
Schooling, Average Years of		0.372***		0.081		0.185***		0.248***
		(12.96)		(1.47)		(4.68)		(6.17)
Inflation, GDP Deflator		-0.013		0.002		-0.003		-0.005
		(-0.71)		(0.08)		(-0.16)		(-0.26)
Investment, % of GDP		-0.212***		-0.116***		-0.151***		-0.109***
		(-11.11)		(-4.05)		(-6.51)		(-4.76)
Island		0.024		0.033		0.030		0.035
		(1.06)		(1.10)		(1.15)		(1.46)
Land Area, Ln		0.280***		0.517***		0.432***		0.335***
		(9.20)		(10.00)		(10.97)		(8.15)
Landlocked		-0.155***		-0.171***		-0.165***		-0.143***
		(-7.72)		(-6.27)		(-7.10)		(-6.72)
Population, Ln		-0.152***		-0.329***		-0.265***		-0.193***
		(-5.05)		(-7.00)		(-7.13)		(-5.02)
Government Consumption, % of GDP		-0.100***		-0.060**		-0.074***		-0.102***
		(-5.30)		(-2.29)		(-3.37)		(-4.91)
Ethno-linguistic Fractionalization		0.060***		0.038		0.046*		0.059**
		(2.67)		(1.24)		(1.76)		(2.34)
Latitude		-0.280***		-0.426***		-0.374***		-0.307***
		(-7.69)		(-8.06)		(-8.60)		(-7.31)
Pacific		-0.005		0.096***		0.059**		0.047*
		(-0.22)		(2.83)		(2.15)		(1.75)
Middle East		0.289***		0.679***		0.539***		0.450***
		(10.88)		(10.80)		(12.74)		(9.72)
Asia		0.117***		0.419***		0.311***		0.219***
		(3.87)		(7.33)		(7.47)		(5.10)
Africa		0.017		0.341***		0.225***		0.164***
		(0.40)		(4.79)		(4.14)		(3.00)
Americas		-0.124***		-0.068		-0.088**		-0.075**
		(-3.94)		(-1.59)		(-2.42)		(-2.17)
N	2863	2332	2863	2332	2863	2332	2683	2250
R^2	0.011	0.285	-0.132	-0.314	-0.000	0.040	-0.018	0.191
F	32.80	54.25	20.39	31.60	0.00190	43.59	49.29	40.63
Underid F Stat P-Value		0.000	0.000	0.000	0.000	0.000	0.000	
Hausman P value			0.000	0.000	0.008	0.000	0.000	0.000

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

Table III-5a
Determinants of Net Exports as a share of GDP, First Stage Results on TOEFL Scores, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.019	0.408***				
			(0.43)	(9.84)				
1100-2200 Hour Dummy			-0.069	0.261***				
			(-1.56)	(6.28)				
2200 Hour and Up Dummy			-0.340***	0.179***				
			(-9.27)	(4.22)				
600 Hour Dummy					-0.028	0.366***		
					(-0.69)	(9.26)		
750 Hour Dummy					0.240***	0.237***		
					(10.65)	(11.14)		
900 Hour Dummy					-0.059**	0.197***		
					(-2.47)	(8.56)		
1100 Hour Dummy					-0.067	0.270***		
					(-1.60)	(6.90)		
1100+ Hour Dummy					-0.031	0.131***		
					(-1.29)	(5.54)		
2200 Hour Dummy					-0.326***	0.247***		
					(-9.33)	(6.01)		
2200+ Hour Dummy					-0.124***	-0.118***		
					(-6.79)	(-6.56)		
Linguistic Proximity-2							0.491***	0.270***
							(29.20)	(12.99)
Covariates			No	Yes	No	Yes	No	Yes
N			2863	2332	2863	2332	2683	2250
R^2			0.109	0.508	0.176	0.539	0.241	0.521
F			116.7	125.9	86.94	117.4	852.4	142.8

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

Table III-6
Determinants of FDI as a share of GDP, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.020	-0.033	-0.085	-0.548***	0.107**	0.029	0.076*	0.003
	(1.07)	(-1.19)	(-1.51)	(-4.28)	(2.30)	(0.36)	(1.87)	(0.03)
Official English		-0.003		0.131***		-0.019		-0.002
		(-0.10)		(3.09)		(-0.58)		(-0.06)
Schooling, Average Years of		0.132***		0.274***		0.115***		0.136***
		(4.06)		(5.60)		(2.99)		(3.15)
Inflation, GDP Deflator		-0.014		-0.022		-0.013		-0.015
		(-0.72)		(-1.01)		(-0.68)		(-0.72)
Investment, % of GDP		0.210***		0.157***		0.216***		0.197***
		(9.80)		(5.99)		(9.54)		(7.86)
Island		-0.034		-0.018		-0.036		-0.038
		(-1.31)		(-0.64)		(-1.39)		(-1.40)
Land Area, Ln		-0.058*		-0.173***		-0.045		-0.037
		(-1.71)		(-3.78)		(-1.18)		(-0.84)
Landlocked		0.018		0.026		0.017		0.025
		(0.80)		(1.07)		(0.76)		(1.12)
Population, Ln		-0.173***		-0.076*		-0.184***		-0.196***
		(-4.99)		(-1.74)		(-4.96)		(-4.53)
Government Consumption, % of GDP		-0.021		-0.026		-0.020		-0.030
		(-0.96)		(-1.10)		(-0.94)		(-1.33)
Ethno-linguistic Fractionalization		0.103***		0.119***		0.101***		0.088***
		(4.05)		(4.33)		(3.97)		(3.24)
Latitude		-0.052		0.027		-0.061		-0.081*
		(-1.25)		(0.57)		(-1.43)		(-1.77)
Pacific		-0.043*		-0.095***		-0.037		-0.045
		(-1.67)		(-3.12)		(-1.38)		(-1.58)
Middle East		-0.094***		-0.279***		-0.072*		-0.081
		(-3.18)		(-5.09)		(-1.82)		(-1.64)
Asia		-0.036		-0.183***		-0.019		-0.030
		(-1.07)		(-3.61)		(-0.47)		(-0.66)
Africa		-0.086*		-0.232***		-0.068		-0.065
		(-1.81)		(-3.75)		(-1.32)		(-1.13)
Americas		-0.060*		-0.074*		-0.058*		-0.065*
		(-1.70)		(-1.96)		(-1.65)		(-1.79)
N	2796	2292	2796	2292	2796	2292	2648	2209
R^2	0.000	0.114	-0.011	-0.022	-0.007	0.112	-0.001	0.108
F	1.153	17.21	2.293	15.92	5.281	17.10	3.489	15.62
Underid F Stat P-Value		0.000	0.000	0.000	0.000	0.000	0.000	
Hausman P value			0.046	0.388	0.0415	1.000	0.123	0.005

Standardized beta coefficients; *t* statistics in parentheses
*** $p < .1$, ** $p < .05$, *** $p < .01$**

Table III-6a
Determinants of FDI as a share of GDP, First Stage Results on TOEFL Scores, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.030	0.411***				
			(0.71)	(10.02)				
1100-2200 Hour Dummy			-0.044	0.265***				
			(-1.03)	(6.42)				
2200 Hour and Up Dummy			-0.338***	0.172***				
			(-9.42)	(4.05)				
600 Hour Dummy					-0.006	0.373***		
					(-0.14)	(9.60)		
750 Hour Dummy					0.218***	0.237***		
					(10.08)	(11.30)		
900 Hour Dummy					-0.050**	0.197***		
					(-2.14)	(8.73)		
1100 Hour Dummy					-0.045	0.274***		
					(-1.09)	(7.09)		
1100+ Hour Dummy					-0.015	0.143***		
					(-0.61)	(6.06)		
2200 Hour Dummy					-0.324***	0.248***		
					(-9.42)	(6.09)		
2200+ Hour Dummy					-0.119***	-0.135***		
					(-6.42)	(-7.47)		
Linguistic Proximity-2							0.480***	0.267***
							(28.15)	(12.82)
Covariates			No	Yes	No	Yes	No	Yes
N			2796	2292	2796	2292	2648	2209
R^2			0.114	0.514	0.167	0.549	0.231	0.526
F			120.3	126.6	79.78	120.2	792.7	143.2

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

Table III-7
Determinants of Emigrant Stock 2000-2008, 2012 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.278***	-0.174	0.613**	-1.717*	0.443**	-0.149	0.348**	-1.989
	(3.27)	(-1.29)	(2.08)	(-1.82)	(2.17)	(-0.57)	(2.12)	(-1.61)
Official English		0.251**		0.590**		0.246**		0.672**
		(2.12)		(2.21)		(2.08)		(2.10)
Schooling, Average Years of		-0.091		0.600		-0.102		0.746
		(-0.65)		(1.30)		(-0.63)		(1.23)
Inflation, GDP Deflator		-0.109		-0.047		-0.110		-0.003
		(-1.12)		(-0.32)		(-1.25)		(-0.02)
Investment, % of GDP		0.046		-0.172		0.049		-0.046
		(0.45)		(-0.86)		(0.50)		(-0.26)
Island		0.219*		0.195		0.219**		0.176
		(1.92)		(1.15)		(2.13)		(0.91)
Land Area, Ln		-0.249*		-0.595**		-0.244*		-0.672*
		(-1.72)		(-2.00)		(-1.73)		(-1.79)
Landlocked		-0.128		-0.015		-0.130		0.029
		(-1.26)		(-0.09)		(-1.40)		(0.16)
Population, Ln		-0.089		0.190		-0.094		0.214
		(-0.59)		(0.68)		(-0.66)		(0.63)
Government Consumption, % of GDP		0.130		0.059		0.131		0.056
		(1.38)		(0.40)		(1.53)		(0.34)
Ethno-linguistic Fractionalization		-0.087		-0.025		-0.088		0.011
		(-0.81)		(-0.16)		(-0.91)		(0.05)
Latitude		0.019		0.383		0.013		0.565
		(0.12)		(1.18)		(0.08)		(1.26)
Pacific		-0.109		-0.159		-0.108		-0.187
		(-0.95)		(-0.93)		(-1.04)		(-0.94)
Middle East		-0.396***		-0.837**		-0.388***		-0.938**
		(-3.07)		(-2.57)		(-2.91)		(-2.14)
Asia		-0.458***		-0.590**		-0.456***		-0.575*
		(-2.77)		(-2.30)		(-3.04)		(-1.92)
Africa		-0.663***		-0.938***		-0.658***		-0.993**
		(-3.03)		(-2.59)		(-3.27)		(-2.26)
Americas		-0.011		0.056		-0.012		0.102
		(-0.07)		(0.23)		(-0.09)		(0.37)
N	129	96	129	96	129	96	119	91
R^2	0.078	0.491	-0.034	-0.363	0.051	0.490	0.076	-0.655
F	10.67	4.420	4.262	1.773	4.626	4.336	4.429	1.372
Underid F Stat P-Value		0.008	0.234	0.002	0.004	0.000	0.076	
Hausman P value			0.236	1.000	0.376	1.000	0.969	1.000

Standardized beta coefficients; *t* statistics in parentheses
*** $p < .1$, ** $p < .05$, *** $p < .01$**

Table III-7a
Determinants of Emigrant Stock in 2012, First Stage Results on TOEFL Scores, 2012 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.027	0.360*				
			(0.14)	(1.86)				
1100-2200 Hour Dummy			-0.072	0.279				
			(-0.40)	(1.53)				
2200 Hour and Up Dummy			-0.303*	0.219				
			(-1.78)	(1.00)				
600 Hour Dummy					-0.016	0.307*		
					(-0.09)	(1.71)		
750 Hour Dummy					0.268**	0.204*		
					(2.53)	(1.86)		
900 Hour Dummy					-0.088	0.153		
					(-0.80)	(1.40)		
1100 Hour Dummy					-0.075	0.236		
					(-0.44)	(1.46)		
1100+ Hour Dummy					-0.012	0.033		
					(-0.11)	(0.31)		
2200 Hour Dummy					-0.292*	0.286		
					(-1.79)	(1.38)		
2200+ Hour Dummy					-0.107	-0.212**		
					(-1.21)	(-2.45)		
Linguistic Proximity-2							0.538***	0.175
							(6.90)	(1.62)
Covariates			No	Yes	No	Yes	No	Yes
N			129	96	129	96	119	91
R^2			0.092	0.658	0.177	0.720	0.290	0.667
F			4.244	7.682	3.705	8.056	47.67	8.595

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

Table III-8
Determinants of Emigrant Stock 2000-2008, 1992 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.037	-0.116	0.295	-0.744	0.217	-0.083	0.283	-0.664
	(0.43)	(-0.98)	(1.41)	(-1.38)	(1.26)	(-0.29)	(1.57)	(-1.57)
Official English		0.002		0.171		-0.007		0.163
		(0.02)		(0.97)		(-0.06)		(1.02)
Schooling, Average Years of		-0.088		0.151		-0.101		0.095
		(-0.68)		(0.63)		(-0.64)		(0.48)
Inflation, GDP Deflator		0.013		0.025		0.012		0.037
		(0.16)		(0.29)		(0.17)		(0.43)
Investment, % of GDP		-0.043		-0.061		-0.042		0.013
		(-0.51)		(-0.68)		(-0.55)		(0.13)
Island		0.201*		0.120		0.205**		0.126
		(1.98)		(0.95)		(2.09)		(1.05)
Land Area, Ln		-0.365**		-0.655**		-0.350**		-0.630***
		(-2.60)		(-2.31)		(-1.96)		(-2.67)
Landlocked		-0.194**		-0.135		-0.197**		-0.142
		(-2.26)		(-1.32)		(-2.41)		(-1.53)
Population, Ln		0.001		0.207		-0.009		0.170
		(0.01)		(0.93)		(-0.06)		(0.92)
Government Consumption, % of GDP		0.122		0.179*		0.119		0.166*
		(1.42)		(1.76)		(1.46)		(1.72)
Ethno-linguistic Fractionalization		-0.050		0.062		-0.056		0.052
		(-0.45)		(0.42)		(-0.50)		(0.38)
Latitude		0.085		0.273		0.075		0.290
		(0.53)		(1.19)		(0.45)		(1.34)
Pacific		-0.030		-0.064		-0.028		-0.054
		(-0.27)		(-0.53)		(-0.28)		(-0.45)
Middle East		-0.371***		-0.582***		-0.360**		-0.576***
		(-3.02)		(-2.67)		(-2.51)		(-2.92)
Asia		-0.456***		-0.599***		-0.448***		-0.594***
		(-3.07)		(-3.05)		(-3.03)		(-3.22)
Africa		-0.531***		-0.645***		-0.525***		-0.582***
		(-2.67)		(-2.81)		(-2.81)		(-2.81)
Americas		-0.060		-0.064		-0.060		-0.041
		(-0.42)		(-0.43)		(-0.46)		(-0.27)
N	137	101	137	101	137	101	127	97
R^2	0.001	0.529	-0.065	0.367	-0.031	0.528	-0.063	0.413
F	0.186	5.477	1.948	4.128	1.572	5.419	2.424	4.193
Underid F Stat P-Value		0.000	0.148	0.000	0.057	0.000	0.004	
Hausman P value			0.178	1.000	0.227	1.000	0.097	1.000

Standardized beta coefficients; *t* statistics in parentheses
*** $p < .1$, ** $p < .05$, *** $p < .01$**

Table III-8a
Determinants of Emigrant Stock in 2012, First Stage Results on TOEFL Scores, 1992 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.037	0.089				
			(0.19)	(0.42)				
1100-2200 Hour Dummy			-0.156	-0.054				
			(-0.82)	(-0.26)				
2200 Hour and Up Dummy			-0.418**	0.176				
			(-2.56)	(0.88)				
600 Hour Dummy					-0.022	0.110		
					(-0.12)	(0.53)		
750 Hour Dummy					0.271***	0.051		
					(2.68)	(0.48)		
900 Hour Dummy					-0.020	0.086		
					(-0.18)	(0.68)		
1100 Hour Dummy					-0.161	-0.025		
					(-0.90)	(-0.13)		
1100+ Hour Dummy					-0.019	-0.003		
					(-0.19)	(-0.03)		
2200 Hour Dummy					-0.406**	0.272		
					(-2.60)	(1.40)		
2200+ Hour Dummy					-0.124	-0.137		
					(-1.51)	(-1.55)		
Linguistic Proximity-2							0.508***	0.279***
							(6.59)	(2.75)
Covariates			No	Yes	No	Yes	No	Yes
N			137	101	137	101	127	97
R^2			0.176	0.612	0.254	0.646	0.258	0.648
F			9.474	6.727	6.269	6.110	43.47	8.558

Standardized beta coefficients; *t* statistics in parentheses
*** *p* < .1, ** *p* < .05, *** *p* < .01**

References

Abadie, Alberto, and Javier Gardeazabal. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* (2003): 113-132.

Acosta Ormaechea, Santiago, and Jiae Yoo. "Tax Composition and Growth: A Broad Cross-Country Perspective." IMF Working Paper No. 12/257, 2012.

Adams Jr, Richard H., and John Page. "Do International Migration and Remittances Reduce Poverty in Developing Countries?" *World Development* 33.10 (2005): 1645-1669.

African Development Bank Group. Domestic Resource Mobilization for Poverty Reduction in East Africa: Burundi Case Study. November 2010.

Akbulut-Yuksel, Mevlude, Hoyt Bleakley, and Aimee Chin. "The Effects of English Proficiency among Childhood Immigrants: Are Hispanics Different?" *Latinos and the Economy* (2011): 255-283.

Ang, James B. "Institutions and the Long-Run Impact of Early Development." *Journal of Development Economics* 105 (2013): 1-18.

Angrist, Joshua D., and Victor Lavy. "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco." *Journal of Labor Economics* (1997): 48-76.

Arcand, J. L., and F. Grin. "Language in Economic Development: Is English Special and is Linguistic Fragmentation Bad?" *English and International Development.* Bristol: Multilingual Matters (2012).

Arnold, Jens Matthias, et al. "Tax Policy for Economic Recovery and Growth." The Economic Journal 121.550 (2011): F59-F80.

Arnold, Jens Matthias. "Do Tax Structures Affect Aggregate Economic Growth?: Empirical Evidence from a Panel of OECD Countries." No. 643. OECD Publishing, 2008.

Azam, Mehtabul, Aimee Chin, and Nishith Prakash. "The Returns to English-Language Skills in India." *Economic Development and Cultural Change* 61.2 (2013): 335-367.

Bambgbose, Ayo. "Language of Instruction Policy and Practice in Africa." Dakar, Senegal: Regional Office for Education in Africa, UNESCO. Retrieved June 5 (2004): 2009.

Barbone, Luca, Richard Bird, and Jaime Vázquez Caro. "The Costs of VAT: A Review of the Literature." Case Network Reports 106 (2012).

Barro, Robert and Jong-Wha Lee, "A New Data Set of Educational Attainment in the World, 1950-2010." Forthcoming, Journal of Development Economics.

Bird, Richard M., and Michael Smart. "Financing Social Expenditures in Developing Countries: Payroll or Value Added Taxes?" No. 1206. International Center for Public Policy, Andrew Young School of Policy Studies, Georgia State University, 2012.

Bird, Richard, J. Martínez-Vázquez, and B. Torgler. "Value-Added Taxes in Developing and Transitional Countries: Lessons and Questions." ITP paper 505 (2005).

Bird, Richard. "Taxation and Development." © World Bank, Washington, DC, 2010.

Birdsall, Nancy. "The New Education Strategy at the World Bank: Time for a Millennium Learning Goal?" Center for Global Development Blog post, May 3rd 2011.

Bleakley, Hoyt, and Aimee Chin. "Language Skills and Earnings: Evidence from Childhood Immigrants." *Review of Economics and Statistics* 86.2 (2004): 481-496.

Bloom, David E., et al. "Geography, demography, and economic growth in Africa." Brookings Papers on Economic Activity 1998.2 (1998): 207-295.

Brock-Utne, Birgit, and Halla B. Holmarsdottir. "Language Policies and Practices in Tanzania and South Africa: Problems and Challenges." *International Journal of Educational Development* 24.1 (2004): 67-83.

Brown, Samuel, and William Gale. "Tax Reform for Growth, Equity, and Revenue." Urban-Brookings Tax Policy Center (2012).

Browning, Martin, and John Burbidge. "Consumption and Income Taxation." Oxford Economic Papers (1990): 281-292.

Bruthiaux, Paul. "Hold your Courses: Language Education, Language Choice, and Economic Development." *TESOL Quarterly* 36.3 (2002): 275-296.

Burgess, Robin, and Nicholas Stern. "Taxation and Development." *Journal of Economic Literature* 31.2 (1993): 762-830.

Cai, Jing, and Ann Harrison. "The Value-Added Tax Reform Puzzle." No. w17532. National Bureau of Economic Research, 2011.

Casale, Daniela, and Dorrit Posel. "English Language Proficiency and Earnings in a Developing Country: The Case of South Africa." *The Journal of Socio-Economics* 40.4 (2011): 385-393.

Chakraborty, Tanika, and Shilpi Kapur. "English language premium: Evidence from a policy experiment in India." (2009). Washington University in St. Louis unpublished paper (2008).

Charalambous, Lorys. “GCC VAT could arrive in Four Years.”Tax-News.com November 29, 2012.

Chiswick, Barry R., and Paul W. Miller. "Language Skills and Earnings among Legalized Aliens." *Journal of Population Economics* 12.1 (1999): 63-89.

Chiswick, Barry R., and Paul W. Miller. "The Endogeneity between Language and Earnings: International Analyses." *Journal of Labor Economics* (1995): 246-288.

Christie, Perry. “A Value Added Tax within a Reformed Tax System.” Government White Paper. February 2013

CIA World Factbook. New York: Skyhorse Publishing, 2012.

Cizek, P., J. Lei, and J. E. Ligthart. “The Determinants of VAT Introduction: A Spatial Duration Analysis.” Center Discussion Paper Series No. 2012-071, 2012.

Cizek, Pavel, Jinghua Lei and Jenny E. Ligthart. “The Determinants of VAT Introduction: A Spatial Duration Analysis.” CentER Discussion Paper Series No. 2012-071, 2012.

Cnossen, S. “Global Trends and Issues in Value Added Taxation.” *International Tax and Public Finance*, 5(3), 1998, 399–428.

Cnossen, Sijbren. "Global Trends and Issues in Value Added Taxation." *International Tax and Public Finance* 5.3 (1998): 399-428.

Coile, Courtney, and Jonathan Gruber. "Future Social Security Entitlements and the Retirement Decision." *The Review of Economics and Statistics* 89.2 (2007): 234-246.

Cox, David R. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society*. Series B (Methodological) (1972): 187-220.

Crystal, David. English as a Global Language. Cambridge University Press, 2003.

Crystal, David. The Cambridge Encyclopedia of the English Language, First edition. Cambridge University Press, 1995.

De Lotbiniere, Max. “Threat to US Jobs Shuts USAid Training.” Guardian Weekly, May 15 2012.

Dehejia, Rajeev H., and Sadek Wahba. "Propensity Score-Matching Methods for Non-Experimental Causal Studies." *Review of Economics and Statistics* 84.1 (2002): 151-161.

Diamond, John W., and George R. Zodrow. "Promoting Growth, Maintaining Progressivity, and Dealing with the Fiscal Crisis CGE Simulations of a Temporary VAT Used for Debt Reduction." *Public Finance Review* 41.6 (2013): 852-884.

Dustmann, Christian, and Arthur Van Soest. "Language and the Earnings of Immigrants." *Industrial and Labor Relations Review* (2002): 473-492

Easterly, William, and Sergio Rebelo. "Fiscal Policy and Economic Growth." *Journal of Monetary Economics* 32.3 (1993): 417-458.

Easterly, William. The Elusive Quest for Growth: Economists Adventures and Misadventure in the Tropics. The MIT press, 2001.

Ebeke, Christian, and Helene Ehrhart. "Tax Revenue Instability in Sub-Saharan Africa: Consequences and Remedies." *Journal of African Economies* 21.1 (2012): 1-27.

Ebrill, L. P., Keen, M., Bodin, J. P., & Summers, V. The Modern VAT. Washington, DC: International Monetary Fund, 2001.

Edmiston, Kelly D., and William F. Fox. 2006. "A Fresh Look at the VAT." In The Challenges of Tax Reform in a Global Economy, edited by James Alm, Jorge Martinez-Vasquez, and Mark Rider, 249-66. New York: Springer.

Education First English Proficiency Index, Executive report, Copyright © 2012 EF Education First Ltd.

Effective Language Learning, http://www.effectivelanguagelearning.com/language-guide/language-difficulty Language Difficulty Ranking, Accessed 2/28/2013

Egger, Peter, and Andrea Lassmann. "The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design." Centre for Economic Policy Research, 2013.

Egger, Peter, and Andrea Lassmann. "The Language Effect in International Trade: A Meta-Analysis." *Economic Letters* 116 (2012) 221-224.

Emran, M. Shahe, and Joseph E. Stiglitz. "On Selective Indirect Tax Reform in Developing Countries." *Journal of Public Economics* 89.4 (2005): 599-623.

Engen, Eric M., and Jonathan Skinner. "Fiscal Policy and Economic Growth." No. w4223. National Bureau of Economic Research, 1992.

Eurostat Europe in Figures Yearbook 2008 on Education .

Fink, Carsten, Aaditya Mattoo, and Ileana Cristina Neagu. "Assessing the Impact of Communication Costs on International Trade." *Journal of International Economics* 67.2 (2005): 428-445.

Fisher, Lloyd D., and D. Y. Lin. "Time-Dependent Covariates in the Cox Proportional-Hazards Regression Model." *Annual Review of Public Health* 20.1 (1999): 145-157.

Furceri, Davide, and Georgios Karras. "Tax Design in the OECD: A Test of the Hines-Summers Hypothesis." *Eastern Economic Journal* 37.2 (2011): 239-247.

Gale, William G., and Samuel Brown. "Tax Reform for Growth, Equity, and Revenue." *Public Finance Review* 41.6 (2013): 721-754.

Gale, William H., and Benjamin H. Harris. "A Value-Added Tax for the United States: Part of the Solution." (2010).

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics* 1.1 (2009): 112-135.

Glick, Reuven, Xueyan Guo, and Michael Hutchison. "Currency Crises, Capital-Account Liberalization, and Selection Bias." *The Review of Economics and Statistics* 88.4 (2006): 698-714.

Gordon, Roger, and Wei Li. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." *Journal of Public Economics* 93.7 (2009): 855-866.

Grandcolas, Christophe. "The Occasional Failure in VAT Implementation: Lessons for the Pacific." *Asia-Pacific Tax Bulletin* (2005): 6-13.

Gremmell, Norman, Richard Kneller, and Ismael Sanz Labrador. "Tax Composition and Economic Growth in OECD Countries." XV Encuentro de Economía Pública: políticas públicas y migración. 2008.

Gupta, Sanjeev, Catherine A. Pattillo, and Smita Wagh. "Effect of Remittances on Poverty and Financial Development in Sub-Saharan Africa." *World Development* 37.1 (2009): 104-115.

Guryan, Jonathan. “Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts.” No. w8269. National Bureau of Economic Research, 2001.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. "Matching as an Econometric Evaluation Estimator." *The Review of Economic Studies* 65.2 (1998): 261-294.

Hejazi, Walid, and Juan Ma. "Gravity, the English Language and International Business." *Multinational Business Review* 19.2 (2011): 152-167.

Hines Jr, James and Mihir A. Desai. "Value-Added Taxes and International Trades: The Evidence." (2005). Unpublished Manuscript.

Hines, J. R. Jr., and L. H. Summers. "How Globalization Affects Tax Design." *Tax Policy and the Economy*, Vol. 23 University of Chicago Press 1427 E 60th St Chicago IL, 2009, 123–57.

Holtz-Eakin, Douglas. "The Case Against VAT." *The VAT Reader, edited by Martin Sullivan* (2011): 96-101.

Hossain, Shahabuddin M. "The Equity Impact of the Value-Added Tax in Bangladesh." Staff Papers-International Monetary Fund (1995): 411-430.

Hoxby, Caroline M. "The Effects of Class Size on Student Achievement: New Evidence from Population Pariation." *The Quarterly Journal of Economics* 115.4 (2000): 1239-1285.

IMF, Financial Data Inquiry 2013

IMF, Islamic Republic of Afghanistan: First Review under Extended Credit Facility. IMF country report No. 12/245 August 2012.

IMF, Kingdom of the Netherlands-Aruba: Concluding Statement of the 2010 Article IV Consultation Mission, 2010

IMF, Republic of Palau: 2012 Article IV Consultation—Staff Report. IMF Country Report No. 12/54 2012.

IMF, Republic of the Marshall Islands: Selected Issues and Statistical Appendix. IMF country report No. 08/186 June 2008.

IMF, World Economic Outlook 2014.

Iorwerth, Aled Ab, and John Whalley. "Efficiency Considerations and the Exemption of Food from Sales and Value Added Taxes." *Canadian Journal of Economics/Revue Canadienne d'Économique* 35.1 (2002): 166-182.

Isphording, Ingo E., and Sebastian Otten. "Linguistic Barriers in the Destination Language Acquisition of Immigrants." *Journal of Economic Behavior & Organization* 105 (2014): 30-50.

Izzak, B and A. Saleh, "Government Wants to Impose VAT." Kuwait Times 01-17-2013

Jepsen, Christopher, and Steven Rivkin. “What is the Tradeoff between Smaller Classes and Teacher Quality?” No. w9205. National Bureau of Economic Research, 2002.

Kapur, Vicky. “VAT May Come to the UAE in 2014: What Will it Mean for You?” November 27, 2012. Emirates 24/7.

Keen, Michael, and Ben Lockwood. "The Value Added Tax: Its Causes and Consequences." *Journal of Development Economics* 92.2 (2010): 138-151.

Keen, Michael, and Ben Lockwood. "The value added tax: Its causes and consequences." *Journal of Development Economics* 92.2 (2010): 138-151.

Keen, Michael. "VAT attacks!" *International Tax and Public Finance* 14.4 (2007): 365-381

Keen, Michael. "VAT, Tariffs, and Withholding: Border Taxes and Informality in Developing Countries." *Journal of Public Economics* 92.10 (2008): 1892-1906.

Keen, Michael. “Taxation and Development Again.” Fiscal Affairs Department IMF: Mimeo, IMF, 2010.

Keen, Michael. “Taxation and Development Again.” Mimeo, IMF, 2010.

Koch, J. "Bringing Improved English teaching to scale: The USAID TELL Program " Paper presented at the annual meeting of the 55th Annual Conference of the Comparative and International Education Society, Fairmont Le Reine, January 10 2014.

Koester, Reinhard B., and Roger C. Kormendi. "Taxation, Aggregate Activity and Economic Growth: Cross-Country Evidence on some Supply-Side Hypotheses." *Economic Inquiry* 27.3 (1989): 367-386.

Kossoudji, Sherrie A. "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men." *Journal of Labor Economics* (1988): 205-228.

Krusell, Per, Vincenzo Quadrini, and Jose-Victor Rios-Rull. "Are Consumption Taxes Really Better than Income Taxes?" *Journal of Monetary Economics* 37.3 (1996): 475-503.

Lane, Phillip R, and Milesi-Ferretti, Gian Maria. “The External Wealth of Nations Mark II: Revised and Extended Estimates of Foreign Assets and Liabilities, 1970–2004.” *Journal of International Economics* 73, November, 223-250 (2007.) Updated to 2011 data.

Lang, Kevin, and Erez Siniver. "The Return to English in a non-English Speaking Country: Russian Immigrants and Native Israelis in Israel." *The BE Journal of Economic Analysis & Policy* 9.1 (2009).

Lee, Chew Ging. "English Language and Economic Growth: Cross-Country Empirical Evidence." *Journal of Economic and Social Studies* 2.1 (2012): 5-20.

Lee, Dongwon, Dongil Kim, and Thomas E. Borcherding. "Tax Structure and Government Spending: Does the Value-Added Tax Increase the Size of Government." *National Tax Journal* 66.3 (2013): 541-70.

Lee, Young, and Roger H. Gordon. "Tax Structure and Economic Growth." *Journal of Public Economics* 89.5 (2005): 1027-1043.

Lu, Bo. "Propensity Score Matching with Time-Dependent Covariates." Biometrics 61.3 (2005): 721-728.

Manage, Neela, and Michael L. Marlow. "The Casual Relation between Federal Expenditures and Receipts." *Southern Economic Journal* 52.3 (1986).

May, Caroline. "USAID Training Foreign Workers for English Speaking Jobs." Daily Caller April 20, 2012.

Melitz, Jacques, and Farid Toubal. "Native Language, Spoken Language, Translation and Trade." *Journal of International Economics* (2014).

Melitz, Jacques. "Language and Foreign Trade." *European Economic Review* 52.4 (2008): 667-699.

Milesi-Ferretti, Gian Maria, and Nouriel Roubini. "Growth Effects of Income and Consumption Taxes." *Journal of Money, Credit and Banking* (1998): 721-744.

Mincer, Jacob. "Schooling, Experience, and Earnings." Human Behavior & Social Institutions No. 2. (1974).

Molnar, Alejandro. "Language Barriers to Foreign Trade: Evidence from Translation Costs." (2013).

Nellor, David. "The Effect of the Value-Added Tax on the Tax Ratio." (1987): 1-28.

Nettle, Daniel. "Linguistic Fragmentation and the Wealth of Nations: The Fishman-Pool Hypothesis Reexamined." *Economic Development and Cultural Change* 48.2 (2000): 335-348.

Nicholson, Michael W. "The Impact of Tax Regimes on International Trade Patterns." *Contemporary Economic Policy* (2012).

Nunan, David. "The Impact of English as a Global Language on Educational Policies and Practices in the Asia-Pacific Region." *Tesol Quarterly* 37.4 (2003): 589-613.

OECD Database on Immigrants in OECD and non-OECD Countries: DIOC, Emigration Rates, http://www.oecd.org/els/mig/dioc.htm, Accessed 2/18/2014

Oh, Chang Hoon, W. Travis Selmier, and Donald Lien. "International Trade, Foreign Direct Investment, and Transaction Costs in Languages." *The Journal of Socio-Economics* 40.6 (2011): 732-735.

Peninsula," Qatar weighs value-added tax proposal." Zawya Connecting Intelligence with Intelligence, January 30 2012.

Politi, Ricardo Batista, and Enlinson Mattos. "Ad-valorem Tax Incidence and After-Tax Price Adjustments: Evidence from Brazilian Basic Basket Food." *Canadian Journal of Economics/Revue canadienne d'économique* 44.4 (2011): 1438-1470.

Posel, Dorrit, and Daniela Casale. "English Language Proficiency and Earnings in a Developing Country: the Case of South Africa." A paper presented at the DPRU/TIPS conference held in Johannesburg from. 2010.

Prah, Kwesi Kwaa. "The Language of Instruction Conundrum in Africa." Meeting on the Implications of Language for Peace and Development (IMPLAN). Retrieved April. Vol. 4. 2008.

Pritchett, Lant. "Where has all the Education Gone?" *The World Bank Economic Review* 15.3 (2001): 367-391.

Rafique, Mahmood. "Bahrain may introduce VAT or Corporate Tax." 24x7 News May 29th 2011.

Rana, Parvaiz Ishfaq, "IMF tells Pakistani Government to Introduce VAT." Asia News Network, 01-24-2013

Rendon, Sílvio. "The Catalan premium: Language and Employment in Catalonia." *Journal of Population Economics* 20.3 (2007): 669-686.

Rooth, Dan-Olof. "The Effect of Language Proficiency on Employment for Immigrants–the Case of Sweden." Submitted for publication (2001.)

Rosenbaum, Paul R., and Donald B. Rubin. "Constructing a Control Group using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *The American Statistician* 39.1 (1985): 33-38.

Rosenbaum, Paul R., and Donald B. Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70.1 (1983): 41-55.

Shastry, Gauri Kartini. "Speaking English in a Globalizing World: Information Technology and Education in India." IZA/World Bank Conference Employment and Development, Bonn. 2007.

Song, Jae Jung. "English as an Official Language in South Korea: Global English or Social Malady?" *Language Problems & Language Planning* 35.1 (2011): 35-55.

Stiglitz, Joseph E. "Development Oriented Tax Policy." Presentation to Congress of International Institute of Public Finance, Prague, 2008.

Stock, J.H. and Yogo, M. 2005. "Testing for Weak Instruments in Linear IV Regression." In D.W.K. Andrews and J.H. Stock, eds. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge: Cambridge University Press, 2005, pp. 80-108. Working paper version: NBER Technical Working Paper 284.

Stockfisch, Jacob A. "Value-Added Taxes and the Size of Government: some Evidence." *National Tax Journal* (1985): 547-552.

Swank, Duane. "Funding the Welfare State: Globalization and the Taxation of Business in Advanced Market Economies." *Political Studies* 46.4 (2002): 671-692.

Toder, Eric, and Joseph Rosenberg. "Effects of Imposing a Value-Added Tax to Replace Payroll Taxes or Corporate Taxes." Tax Policy Center Publication. April 7 (2010).

Treisman, Daniel. "Decentralization and the Quality of Government." Manuscript UCLA (2002).

Ufier, Alex. "Quasi-Experimental Analysis on the Effects of Adoption of a Value Added Tax." *Economic Inquiry* 52.4 (2014): 1364-1379.

University of Pennsylvania, Penn World Tables 2013.

USAID, "English for Engineers in Afghanistan" USAID Leadership Success stories, December 06, 2013

USAID, "Rwandan Teachers Switch to English Instruction" USAID Leadership Success stories, March 06, 2014

Van den Noord, Paul, and Christopher Heady. "Tax Design, Economic Efficiency and Growth." *Economic Efficiency and Growth* (March 21, 2002) (2012).

Vartia, Laura. How do taxes affect investment and productivity?: An industry-level analysis of OECD countries. No. 656. OECD Publishing, 2008.

Von Furstenberg, George M., R. Jeffery Green, and Jin-Ho Jeong. "Tax and Spend, or Spend and Tax?" *The Review of Economics and Statistics* (1986): 179-188.

Warschauer, Mark. "The Changing Global Economy and the Future of English Teaching." *Tesol Quarterly* 34.3 (2000): 511-535.

Wedell, Martin. "Developing a capacity to make “English for Everyone: worthwhile: Reconsidering Outcomes and how to Start Achieving them." *International Journal of Educational Development* 28.6 (2008): 628-639.

Wolhuter, C. C., and S. C. Steyn. "Learning from south-south comparison: the education systems of South Africa and Madagascar." *South African Journal of Education* 23.1 (2006): 29-35.

World Development Indicators, http://data.worldbank.org/data-catalog/world-development-indicators, The World Bank.

Zagler, Martin, and Georg Dürnecker. "Fiscal Policy and Economic Growth." *Journal of Economic Surveys* 17.3 (2003): 397-418.

Appendices

Appendices are labelled with a roman numeral corresponding to their chapter.

Appendix I-A: Countries Included

1. *Countries Included in Specification (1) of Table I-2, 112 countries (Most Restrictive):*

Albania; Algeria; Angola; Australia; Bahamas, The; Bahrain; Bangladesh; Barbados; Belize; Benin; Bhutan; Bosnia and Herzegovina; Botswana; Brunei Darussalam; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Republic; Chad; China; Comoros; Congo, Dem. Rep.; Congo, Rep.; Croatia; Cyprus; Djibouti; Egypt, Arab Rep.; El Salvador; Equatorial Guinea; Eritrea; Ethiopia; Fiji; Finland; Gabon; Gambia, The; Ghana; Grenada; Guinea; Guinea-Bissau; Guyana; Hungary; Iceland; India; Iran, Islamic Rep.; Japan; Jordan; Kenya; Kuwait; Lao PDR; Lebanon; Lesotho; Liberia; Libya; Macedonia, FYR; Madagascar; Malawi; Malaysia; Maldives; Mali; Malta; Mauritania; Mauritius; Mongolia; Morocco; Mozambique; Namibia; Nepal; New Zealand; Niger; Oman; Pakistan; Papua New Guinea; Paraguay; Philippines; Poland; Portugal; Romania; Rwanda; Saudi Arabia; Serbia; Sierra Leone; Singapore; Slovenia; Solomon Islands; South Africa; Spain; Sri Lanka; St. Lucia; St. Vincent and the Grenadines; Sudan; Suriname; Swaziland; Syrian Arab Republic; Tanzania; Thailand; Togo; Tonga; Trinidad and Tobago; Tunisia; Uganda; United Arab Emirates; United States; Vanuatu; Venezuela, RB; Vietnam; Yemen, Rep.; Zambia; Zimbabwe.

2. *Countries Included in Specification (5) of Table I-2, 192 countries (Least Restrictive):*

Afghanistan; Albania; Algeria; Andorra; Angola; Antigua and Barbuda; Argentina; Armenia; Australia; Austria; Azerbaijan; Bahamas, The; Bahrain; Bangladesh; Barbados; Belarus; Belize; Benin; Bhutan; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Brunei Darussalam; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Republic; Chad; Chile; China; Colombia; Comoros; Congo, Dem. Rep.; Congo, Rep.; Costa Rica; Croatia; Cuba; Cyprus; Czech Republic; Denmark; Djibouti; Dominica; Dominican Republic; Ecuador; Egypt, Arab Rep.; El Salvador; Equatorial Guinea; Eritrea; Estonia; Ethiopia; Fiji; Finland; France; Gabon; Gambia, The; Georgia; Germany; Ghana; Greece; Grenada; Guatemala; Guinea; Guinea-Bissau; Guyana; Haiti; Honduras; Hong Kong SAR, China; Hungary; Iceland; India; Indonesia; Iran, Islamic Rep.; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Kiribati; Korea, Dem. Rep.; Korea, Rep.; Kosovo; Kuwait; Kyrgyz Republic; Lao PDR; Latvia; Lebanon; Lesotho; Liberia; Libya; Liechtenstein; Lithuania; Macao SAR, China; Macedonia, FYR; Madagascar; Malawi; Malaysia; Maldives; Mali; Malta; Marshall Islands; Mauritania; Mauritius; Mexico; Micronesia, Fed. Sts.; Moldova; Monaco; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; Norway; Oman; Pakistan; Palau; Panama; Papua New Guinea; Paraguay; Peru;

Philippines; Poland; Portugal; Qatar; Romania; Russian Federation; Rwanda; Samoa; San Marino; Sao Tome and Principe; Saudi Arabia; Senegal; Serbia; Seychelles; Sierra Leone; Singapore; Slovak Republic; Slovenia; Solomon Islands; Somalia; South Africa; Spain; Sri Lanka; St. Kitts and Nevis; St. Lucia; St. Vincent and the Grenadines; Sudan; Suriname; Swaziland; Sweden; Switzerland; Syrian Arab Republic; Tajikistan; Tanzania; Thailand; Timor-Leste; Togo; Tonga; Trinidad and Tobago; Tunisia; Turkey; Turkmenistan; Tuvalu; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States; Uruguay; Uzbekistan; Vanuatu; Venezuela, RB; Vietnam; West Bank and Gaza; Yemen, Rep.; Zambia; Zimbabwe

3. Countries included in Matching

Note the countries included in each matching specifications will be slightly different in each case, as the method here demands 13 continuous years of data (7 before Year of Adoption and 5 after plus the actual year of the adoption.) Thus, one can only study the impact of adoptions from 1986-2007, even though one can study the determination of adoptions up through 2012. Exact countries included in each specification are available upon request from the author.

Appendix I-B: Data Definitions

Agriculture: Percent of GDP contributed by agricultural activities, from the World Development Indicators.

Alliance of Small Island States: Equal to 1 if a country identifies as a member of the Alliance of Small Island States at any time in the sample period, from the Alliance of Small Island States.

Bordering Countries with VAT: Number of bordering countries with a VAT divided by the number of bordering countries, from author's calculations.

British Commonwealth: Equal to 1 if a country is a member of the British Commonwealth at any time in the sample period, from the Commonwealth Network.

Exports: Exports as a percent of GDP, from the World Development Indicators.

Federal: Equal to 1 if a country is governed as a federation, from Treisman (2002) with updates from the author.

Former Soviet Republic: Equal to 1 if a country was created by the collapse of a Warsaw Pact Nation, from author's calculations.

Former Warsaw Pact: Equal to 1 if a country was a Warsaw Pact Nation, from author's calculations.

Francophonie: Equal to 1 if a country is a member of the Francophonie at any time in the sample period, from Organisation internationale de la Francophonie.

GDP per capita growth: Real GDP per capita growth, from the World Development Indicators

GDP Per Capita, Log: Natural log of GDP per capita as measured by current US dollars, from the World Development Indicators.

GDP, Log: Natural log of GDP as measured by current US dollars, from the World Development Indicators.

Government Consumption: Government Consumption, as a percent of GDP from the Penn World Tables.

IMF Lending Dummy: Equal to 1 if the country has a nonzero balance in either a PRG or GRA credit facility in any months of that year, a stock measure, from the IMF Financial Data query tool.

Imports: Imports as a percent of GDP, from the World Development Indicators.

Inflation: Yearly inflation as measured by the CPI, from the World Development Indicators.

Investment: Gross capital formation as a percent of GDP from the world development indicators.

Island: Equal to 1 if a country has no other countries it borders by land, from author's calculations.

Landlocked: Equal to 1 if a country has no land bordering an ocean, from author's calculations.

Natural Resource Rents: Percent of GDP coming from natural resource production, from the World Development Indicators.

Net Exports: Exports as a percent of GDP minus imports as a percent of GDP, from the World Development Indicators.

Population Growth: Growth rate of the population, from the World Development Indicators.

Population, Log: Natural log of population, from the World Development Indicators.

% of Population 0-14: Percent of population between 0 and 14 years of age, from the World Development Indicators.

% of Population 65+: Percent of population above 65 years of age, from the World Development Indicators.

Size, Log Square Kilometers: Natural log of the surface area of a country, from the World Development Indicators.

Trade: Exports as a percent of GDP plus imports as a share of GDP, from the World Development Indicators.

Table I-6
Comparison of Matching Methods, Warsaw Pact Omitted

Variable, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	1 Nearest Neighbor, No Replacement	3 Nearest Neighbors	Radius, .025	Radius, .05	Kernel, Epanechnikov	Kernel, Normal	Total Significant at 10% level
Investment	1.77*	1.77*	1.34	1.38	2.01**	1.94*	4
GDP per Capita Growth	1.68*	0.74	0.44	0.72	-.03	0.83	1
Trade	0.08	0.52	-0.58	-0.42	1.22	1.08	0
Inflation	-1.35	-1.05	-0.95	-1.77*	-1.79*	-1.90*	3
Government Consumption	-1.88*	-1.61	-1.80*	-1.85*	-1.08	-1.18	3

Variable, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	1 Nearest Neighbor, No Replacement	3 Nearest Neighbors	Radius, .025	Radius, .05	Kernel, Epanechnikov	Kernel, Normal	Total Significant at 10% level
Investment	-0.91	-0.80	-0.87	-0.85	-0.91	-0.74	0
GDP per Capita Growth	-0.42	-0.08	-0.03	-.07	-0.10	-0.19	0
Trade	0.26	0.15	0.37	0.24	0.46	0.57	0
Inflation	0.72	0.66	1.25	1.20	0.91	0.11	0
Government Consumption	-1.18	-1.02	-1.17	-1.10	-1.22	-1.03	0

***t* statistics in parentheses.**
* $p < .1$, ** $p < .05$, *** $p < .01$

***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-7
Impact of VAT adoption on Investment, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Investment, 7 Years Pre-Adoption	Unmatched	23.27	23.86	-0.60	1.24	-0.48
	ATT	23.27	22.98	0.28	2.08	0.14
Investment, 6 Years Pre-Adoption	Unmatched	22.88	23.73	-0.85	1.24	-0.69
	ATT	22.88	23.94	-1.07	2.07	-0.52
Investment, 5 Years Pre-Adoption	Unmatched	22.53	23.59	-1.06	1.25	-0.84
	ATT	22.53	23.24	-0.71	1.98	-0.36
Investment, 4 Years Pre-Adoption	Unmatched	22.80	23.27	-0.47	1.26	-0.38
	ATT	22.80	22.06	0.73	1.79	0.41
Investment, 3 Years Pre-Adoption	Unmatched	21.95	23.03	-1.08	1.27	-0.85
	ATT	21.95	22.42	-0.47	1.75	-0.27
Investment, 2 Years Pre-Adoption	Unmatched	21.50	22.96	-1.45	1.28	-1.14
	ATT	21.50	22.56	-1.05	1.92	-0.55
Investment, 1 Year Pre-Adoption	Unmatched	22.44	22.85	-0.41	1.30	-0.32
	ATT	22.44	22.44	-0.01	1.91	0.00
Investment, Year of Adoption	Unmatched	21.66	22.78	-1.12	1.30	-0.86
	ATT	21.66	22.32	-0.66	1.79	-0.37
Investment, 1 Year Post-Adoption	Unmatched	22.32	22.67	-0.34	1.30	-0.26
	ATT	22.32	21.85	0.48	1.52	0.31
Investment, 2 Years Post-Adoption	Unmatched	22.68	22.70	-0.03	1.31	-0.02
	ATT	22.68	21.53	1.15	1.67	0.69
Investment, 3 Years Post-Adoption	Unmatched	22.91	22.67	0.24	1.30	0.18
	ATT	22.91	21.20	1.71	1.54	1.11
Investment, 4 Years Post-Adoption	Unmatched	22.53	22.74	-0.21	1.29	-0.16
	ATT	22.53	21.11	1.43	1.48	0.97
Investment, 5 Years Post-Adoption	Unmatched	22.69	22.80	-0.12	1.29	-0.09
	ATT	22.69	20.71	1.98	1.64	1.21
Investment, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**1.18**	**-0.15**	**1.33**	**1.06**	**1.26**
	ATT	**1.18**	**-1.85**	**3.03**	**1.80**	**1.69***
Investment, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**-1.77**	**-0.91**	**-0.86**	**1.06**	**-0.81**
	ATT	**-1.77**	**-0.43**	**-1.34**	**1.70**	**-0.79**

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$
Treated Observations: 77
Untreated Observations: 639
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-8
Impact of VAT adoption on GDP per Capita Growth, Single Nearest Neighbor

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
GDP per Capita Growth, 7 Years Pre-Adoption	Unmatched	1.69	1.21	0.48	0.91	0.53
	ATT	1.69	1.64	0.04	1.45	0.03
GDP per Capita Growth, 6 Years Pre-Adoption	Unmatched	1.99	1.17	0.82	0.83	0.98
	ATT	1.99	1.42	0.57	1.11	0.51
GDP per Capita Growth, 5 Years Pre-Adoption	Unmatched	1.30	1.24	0.06	0.83	0.07
	ATT	1.30	2.58	-1.29	1.06	-1.22
GDP per Capita Growth, 4 Years Pre-Adoption	Unmatched	1.18	1.06	0.12	0.66	0.18
	ATT	1.18	2.57	-1.39	0.96	-1.45
GDP per Capita Growth, 3 Years Pre-Adoption	Unmatched	1.53	1.12	0.41	0.71	0.58
	ATT	1.53	1.13	0.41	1.26	0.32
GDP per Capita Growth, 2 Years Pre-Adoption	Unmatched	1.49	1.44	0.05	0.68	0.08
	ATT	1.49	0.79	0.70	1.05	0.67
GDP per Capita Growth, 1 Year Pre-Adoption	Unmatched	1.65	1.44	0.21	0.77	0.27
	ATT	1.65	1.05	0.60	1.02	0.59
GDP per Capita Growth, Year of Adoption	Unmatched	2.63	1.60	1.03	0.78	1.32
	ATT	2.63	0.36	2.26	1.04	2.17**
GDP per Capita Growth, 1 Year Post-Adoption	Unmatched	1.69	1.76	-0.08	0.77	-0.10
	ATT	1.69	1.33	0.35	0.93	0.38
GDP per Capita Growth, 2 Years Post-Adoption	Unmatched	2.32	1.92	0.40	0.74	0.55
	ATT	2.32	1.31	1.01	0.84	1.20
GDP per Capita Growth, 3 Years Post-Adoption	Unmatched	2.08	1.84	0.24	0.77	0.31
	ATT	2.08	0.50	1.58	0.89	1.79*
GDP per Capita Growth, 4 Years Post-Adoption	Unmatched	2.33	1.80	0.53	0.77	0.70
	ATT	2.33	0.36	1.97	1.02	1.93*
GDP per Capita Growth, 5 Years Post-Adoption	Unmatched	2.07	1.70	0.37	0.70	0.53
	ATT	2.07	0.72	1.34	0.97	1.38
GDP per Capita Growth, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**0.57**	**0.26**	**0.32**	**0.92**	**0.34**
	ATT	**0.57**	**-0.07**	**0.64**	**1.32**	**0.49**
GDP per Capita Growth, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**-0.19**	**0.23**	**-0.43**	**1.06**	**-0.40**
	ATT	**-0.19**	**-0.85**	**0.66**	**1.76**	**0.37**

*** $p < .1$, ** $p < .05$, *** $p < .01$**
Treated Observations: 78
Untreated Observations: 683
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-9
Impact of VAT adoption on Trade, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Trade, 7 Years Pre-Adoption	Unmatched	71.99	84.63	-12.63	6.34	-1.99**
	ATT	71.99	63.59	8.41	9.65	0.87
Trade, 6 Years Pre-Adoption	Unmatched	69.49	84.50	-15.02	6.36	-2.36**
	ATT	69.49	64.48	5.00	10.33	0.48
Trade, 5 Years Pre-Adoption	Unmatched	70.59	83.70	-13.11	6.28	-2.09**
	ATT	70.59	65.05	5.54	10.70	0.52
Trade, 4 Years Pre-Adoption	Unmatched	69.93	83.12	-13.20	6.18	-2.13**
	ATT	69.93	61.91	8.01	9.66	0.83
Trade, 3 Years Pre-Adoption	Unmatched	70.79	82.78	-12.00	6.12	-1.96*
	ATT	70.79	61.90	8.89	9.33	0.95
Trade, 2 Years Pre-Adoption	Unmatched	71.29	82.81	-11.52	6.08	-1.90*
	ATT	71.29	61.39	9.90	9.21	1.07
Trade, 1 Year Pre-Adoption	Unmatched	73.47	82.98	-9.51	6.10	-1.56
	ATT	73.47	64.88	8.59	9.94	0.86
Trade, Year of Adoption	Unmatched	73.91	82.85	-8.94	6.13	-1.46
	ATT	73.91	64.89	9.02	10.24	0.88
Trade, 1 Year Post-Adoption	Unmatched	76.30	83.50	-7.21	6.18	-1.17
	ATT	76.30	65.13	11.16	10.11	1.10
Trade, 2 Years Post-Adoption	Unmatched	77.91	84.23	-6.32	6.19	-1.02
	ATT	77.91	67.19	10.71	9.79	1.09
Trade, 3 Years Post-Adoption	Unmatched	78.30	84.67	-6.38	6.12	-1.04
	ATT	78.30	67.03	11.27	9.70	1.16
Trade, 4 Years Post-Adoption	Unmatched	78.22	85.10	-6.88	6.08	-1.13
	ATT	78.22	70.36	7.86	9.85	0.80
Trade, 5 Years Post-Adoption	Unmatched	80.89	85.71	-4.83	6.08	-0.79
	ATT	80.89	72.06	8.83	10.15	0.87
Trade, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**9.60**	**2.91**	**6.69**	**3.19**	**2.10****
	ATT	**9.60**	**10.67**	**-1.07**	**4.21**	**-0.25**
Trade, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**-0.71**	**-1.82**	**1.12**	**2.79**	**0.40**
	ATT	**-0.71**	**-2.20**	**1.49**	**4.28**	**0.35**

*** $p < .1$, ** $p < .05$, *** $p < .01$**
Treated Observations: 76
Untreated Observations: 662
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-10
Impact of VAT adoption on Inflation, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Inflation, 7 Years Pre-Adoption	Unmatched	22.56	82.81	-60.26	120.81	-0.50
	ATT	22.56	26.47	-3.91	12.09	-0.32
Inflation, 6 Years Pre-Adoption	Unmatched	20.17	89.57	-69.40	121.88	-0.57
	ATT	20.17	29.56	-9.39	12.26	-0.77
Inflation, 5 Years Pre-Adoption	Unmatched	23.81	90.06	-66.25	121.91	-0.54
	ATT	23.81	28.94	-5.13	14.36	-0.36
Inflation, 4 Years Pre-Adoption	Unmatched	22.22	88.58	-66.36	121.78	-0.54
	ATT	22.22	18.04	4.18	12.05	0.35
Inflation, 3 Years Pre-Adoption	Unmatched	16.73	81.84	-65.12	120.50	-0.54
	ATT	16.73	14.35	2.37	9.58	0.25
Inflation, 2 Years Pre-Adoption	Unmatched	12.38	82.15	-69.77	120.57	-0.58
	ATT	12.38	10.31	2.07	4.27	0.48
Inflation, 1 Year Pre-Adoption	Unmatched	13.23	76.61	-63.38	118.60	-0.53
	ATT	13.23	13.21	0.02	5.14	0.00
Inflation, Year of Adoption	Unmatched	14.13	84.30	-70.17	120.82	-0.58
	ATT	14.13	12.30	1.83	4.58	0.40
Inflation, 1 Year Post-Adoption	Unmatched	11.28	83.47	-72.19	120.71	-0.60
	ATT	11.28	13.90	-2.62	3.83	-0.68
Inflation, 2 Years Post-Adoption	Unmatched	11.37	79.74	-68.38	120.37	-0.57
	ATT	11.37	13.81	-2.44	4.08	-0.60
Inflation, 3 Years Post-Adoption	Unmatched	22.95	76.75	-53.80	120.20	-0.45
	ATT	22.95	95.35	-72.40	47.98	-1.51
Inflation, 4 Years Post-Adoption	Unmatched	10.10	68.35	-58.25	117.90	-0.49
	ATT	10.10	166.67	-156.57	85.82	-1.82*
Inflation, 5 Years Post-Adoption	Unmatched	7.24	67.80	-60.56	117.90	-0.51
	ATT	7.24	73.55	-66.31	35.06	-1.89*
Inflation, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**-5.14**	**-14.35**	**9.21**	**167.14**	**0.06**
	ATT	**-5.14**	**63.24**	**-68.38**	**34.72**	**-1.97***
Inflation, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**-10.18**	**-0.67**	**-9.51**	**168.77**	**-0.06**
	ATT	**-10.18**	**-16.16**	**5.98**	**12.64**	**0.47**

*** $p < .1$, ** $p < .05$, *** $p < .01$**
Treated Observations: 78
Untreated Observations: 684
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-11
Impact of VAT adoption on Government Consumption, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Government Consumption, 7 Years Pre-Adoption	Unmatched	11.70	15.55	-3.86	1.30	-2.97***
	ATT	11.70	10.20	1.50	1.73	0.86
Government Consumption, 6 Years Pre-Adoption	Unmatched	11.48	15.53	-4.05	1.29	-3.14***
	ATT	11.48	10.34	1.14	1.70	0.67
Government Consumption, 5 Years Pre-Adoption	Unmatched	11.55	15.47	-3.92	1.30	-3.03***
	ATT	11.55	10.25	1.30	1.64	0.79
Government Consumption, 4 Years Pre-Adoption	Unmatched	11.70	15.40	-3.70	1.29	-2.87***
	ATT	11.70	10.12	1.58	1.60	0.99
Government Consumption, 3 Years Pre-Adoption	Unmatched	11.44	15.23	-3.79	1.27	-2.99***
	ATT	11.44	10.12	1.32	1.59	0.83
Government Consumption, 2 Years Pre-Adoption	Unmatched	11.34	15.08	-3.74	1.26	-2.98***
	ATT	11.34	9.99	1.35	1.46	0.92
Government Consumption, 1 Year Pre-Adoption	Unmatched	11.16	14.94	-3.78	1.25	-3.03***
	ATT	11.16	10.23	0.93	1.62	0.57
Government Consumption, Year of Adoption	Unmatched	10.77	14.69	-3.92	1.22	-3.23***
	ATT	10.77	10.46	0.31	1.66	0.18
Government Consumption, 1 Year Post-Adoption	Unmatched	10.80	14.49	-3.69	1.18	-3.12***
	ATT	10.80	9.95	0.84	1.45	0.58
Government Consumption, 2 Years Post-Adoption	Unmatched	10.72	14.26	-3.54	1.14	-3.12***
	ATT	10.72	10.01	0.71	1.51	0.47
Government Consumption, 3 Years Post-Adoption	Unmatched	10.71	14.14	-3.43	1.12	-3.07***
	ATT	10.71	9.74	0.97	1.48	0.65
Government Consumption, 4 Years Post-Adoption	Unmatched	10.51	14.11	-3.60	1.10	-3.28***
	ATT	10.51	10.02	0.48	1.38	0.35
Government Consumption, 5 Years Post-Adoption	Unmatched	10.27	14.05	-3.78	1.07	-3.52***
	ATT	10.27	10.26	0.01	1.47	0.01
Government Consumption, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**-1.07**	**-1.03**	**-0.04**	**0.69**	**-0.05**
	ATT	**-1.07**	**0.27**	**-1.34**	**0.77**	**-1.75***
Government Consumption, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**-0.36**	**-0.47**	**0.11**	**0.60**	**0.19**
	ATT	**-0.36**	**-0.21**	**-0.15**	**0.70**	**-0.21**

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$
Treated Observations: 80
Untreated Observations: 716
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-12
Impact of VAT adoption on Imports, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Imports, 7 Years Pre-Adoption	Unmatched	40.66	48.62	-7.96	3.59	-2.21**
	ATT	40.66	34.42	6.24	5.43	1.15
Imports, 6 Years Pre-Adoption	Unmatched	39.04	48.38	-9.35	3.56	-2.63***
	ATT	39.04	35.19	3.85	5.89	0.65
Imports, 5 Years Pre-Adoption	Unmatched	39.42	47.78	-8.36	3.52	-2.38**
	ATT	39.42	35.18	4.24	5.99	0.71
Imports, 4 Years Pre-Adoption	Unmatched	39.26	47.06	-7.80	3.46	-2.25**
	ATT	39.26	33.05	6.22	5.41	1.15
Imports, 3 Years Pre-Adoption	Unmatched	39.49	46.57	-7.09	3.40	-2.08**
	ATT	39.49	33.60	5.88	5.15	1.14
Imports, 2 Years Pre-Adoption	Unmatched	39.44	46.36	-6.92	3.37	-2.05**
	ATT	39.44	33.51	5.93	5.11	1.16
Imports, 1 Year Pre-Adoption	Unmatched	40.60	46.21	-5.61	3.37	-1.67*
	ATT	40.60	35.40	5.20	5.51	0.94
Imports, Year of Adoption	Unmatched	40.34	46.03	-5.69	3.36	-1.69*
	ATT	40.34	35.18	5.15	5.66	0.91
Imports, 1 Year Post-Adoption	Unmatched	41.87	46.18	-4.31	3.40	-1.27
	ATT	41.87	34.59	7.28	5.63	1.29
Imports, 2 Years Post-Adoption	Unmatched	43.31	46.38	-3.07	3.42	-0.90
	ATT	43.31	36.05	7.26	5.38	1.35
Imports, 3 Years Post-Adoption	Unmatched	43.08	46.54	-3.46	3.36	-1.03
	ATT	43.08	35.57	7.51	5.19	1.45
Imports, 4 Years Post-Adoption	Unmatched	42.78	46.72	-3.94	3.33	-1.18
	ATT	42.78	36.99	5.79	5.26	1.10
Imports, 5 Years Post-Adoption	Unmatched	44.13	46.94	-2.81	3.29	-0.86
	ATT	44.13	37.96	6.17	5.46	1.13
Imports, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**4.69**	**0.58**	**4.11**	**1.88**	**2.18****
	ATT	**4.69**	**4.44**	**0.24**	**2.36**	**0.10**
Imports, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**-1.22**	**-2.25**	**1.04**	**1.74**	**0.59**
	ATT	**-1.22**	**-0.90**	**-0.31**	**2.61**	**-0.12**

*** $p < .1$, ** $p < .05$, *** $p < .01$**
Treated Observations: 76
Untreated Observations: 662
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

Table I-13
Impact of VAT adoption on Exports, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Exports, 7 Years Pre-Adoption	Unmatched	31.33	36.01	-4.68	3.21	-1.46
	ATT	31.33	29.17	2.16	4.81	0.45
Exports, 6 Years Pre-Adoption	Unmatched	30.45	36.12	-5.67	3.25	-1.75*
	ATT	30.45	29.29	1.16	5.00	0.23
Exports, 5 Years Pre-Adoption	Unmatched	31.18	35.93	-4.75	3.21	-1.48
	ATT	31.18	29.87	1.31	5.18	0.25
Exports, 4 Years Pre-Adoption	Unmatched	30.66	36.06	-5.40	3.19	-1.69*
	ATT	30.66	28.87	1.79	4.75	0.38
Exports, 3 Years Pre-Adoption	Unmatched	31.30	36.21	-4.91	3.17	-1.55
	ATT	31.30	28.30	3.01	4.62	0.65
Exports, 2 Years Pre-Adoption	Unmatched	31.85	36.44	-4.60	3.16	-1.46
	ATT	31.85	27.88	3.96	4.53	0.88
Exports, 1 Year Pre-Adoption	Unmatched	32.87	36.77	-3.90	3.19	-1.22
	ATT	32.87	29.48	3.38	4.83	0.70
Exports, Year of Adoption	Unmatched	33.57	36.82	-3.25	3.23	-1.01
	ATT	33.57	29.70	3.87	4.97	0.78
Exports, 1 Year Post-Adoption	Unmatched	34.42	37.33	-2.90	3.26	-0.89
	ATT	34.42	30.54	3.88	4.98	0.78
Exports, 2 Years Post-Adoption	Unmatched	34.60	37.85	-3.25	3.26	-1.00
	ATT	34.60	31.14	3.46	4.87	0.71
Exports, 3 Years Post-Adoption	Unmatched	35.22	38.13	-2.91	3.23	-0.90
	ATT	35.22	31.45	3.76	4.90	0.77
Exports, 4 Years Post-Adoption	Unmatched	35.44	38.38	-2.94	3.21	-0.92
	ATT	35.44	33.37	2.07	4.95	0.42
Exports, 5 Years Post-Adoption	Unmatched	36.76	38.77	-2.01	3.23	-0.62
	ATT	36.76	34.10	2.65	5.05	0.53
Exports, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**4.91**	**2.33**	**2.59**	**1.63**	**1.59**
	ATT	**4.91**	**6.22**	**-1.31**	**2.26**	**-0.58**
Exports, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**0.51**	**0.43**	**0.08**	**1.44**	**0.06**
	ATT	**0.51**	**-1.29**	**1.80**	**2.13**	**0.85**

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$
Treated Observations: 76
Untreated Observations: 662
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

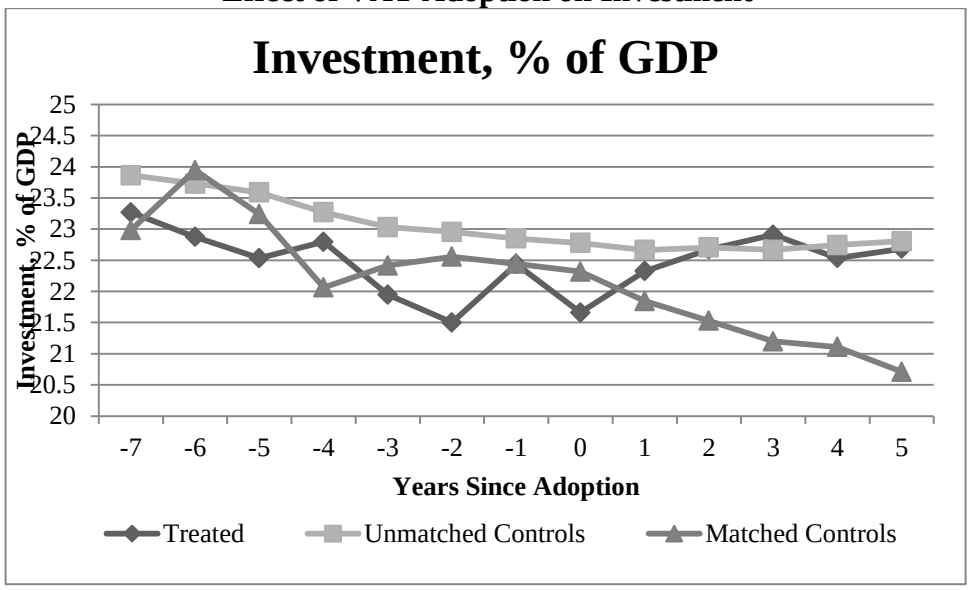
Table I-14
Impact of VAT adoption on Net Exports, Single Nearest Neighbor Matching

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Net Exports, 7 Years Pre-Adoption	Unmatched	-9.33	-12.60	3.28	2.50	1.31
	ATT	-9.33	-5.24	-4.08	3.48	-1.17
Net Exports, 6 Years Pre-Adoption	Unmatched	-8.59	-12.27	3.68	2.43	1.51
	ATT	-8.59	-5.90	-2.69	3.54	-0.76
Net Exports, 5 Years Pre-Adoption	Unmatched	-8.24	-11.85	3.61	2.44	1.48
	ATT	-8.24	-5.31	-2.93	3.32	-0.88
Net Exports, 4 Years Pre-Adoption	Unmatched	-8.60	-11.01	2.41	2.46	0.98
	ATT	-8.60	-4.18	-4.42	3.19	-1.39
Net Exports, 3 Years Pre-Adoption	Unmatched	-8.18	-10.36	2.18	2.42	0.90
	ATT	-8.18	-5.31	-2.88	2.94	-0.98
Net Exports, 2 Years Pre-Adoption	Unmatched	-7.60	-9.92	2.32	2.39	0.97
	ATT	-7.60	-5.63	-1.97	2.88	-0.68
Net Exports, 1 Year Pre-Adoption	Unmatched	-7.73	-9.44	1.71	2.44	0.70
	ATT	-7.73	-5.91	-1.82	2.95	-0.62
Net Exports, Year of Adoption	Unmatched	-6.76	-9.21	2.45	2.43	1.01
	ATT	-6.76	-5.48	-1.29	2.94	-0.44
Net Exports, 1 Year Post-Adoption	Unmatched	-7.45	-8.85	1.40	2.46	0.57
	ATT	-7.45	-4.04	-3.40	3.26	-1.04
Net Exports, 2 Years Post-Adoption	Unmatched	-8.71	-8.53	-0.18	2.52	-0.07
	ATT	-8.71	-4.91	-3.80	3.10	-1.23
Net Exports, 3 Years Post-Adoption	Unmatched	-7.86	-8.41	0.55	2.45	0.22
	ATT	-7.86	-4.12	-3.74	2.80	-1.34
Net Exports, 4 Years Post-Adoption	Unmatched	-7.35	-8.34	1.00	2.42	0.41
	ATT	-7.35	-3.62	-3.72	2.71	-1.37
Net Exports, 5 Years Post-Adoption	Unmatched	-7.37	-8.17	0.80	2.36	0.34
	ATT	-7.37	-3.85	-3.52	2.77	-1.27
Net Exports, Change 2 Years Pre-Adoption to 5 Years Post-Adoption	**Unmatched**	**0.22**	**1.75**	**-1.52**	**1.49**	**-1.02**
	ATT	**0.22**	**1.78**	**-1.55**	**1.90**	**-0.82**
Net Exports, Change 7 Years Pre-Adoption to 2 Years Pre-Adoption	**Unmatched**	**1.73**	**2.68**	**-0.96**	**1.56**	**-0.61**
	ATT	**1.73**	**-0.39**	**2.12**	**2.08**	**1.02**

$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$
Treated Observations: 76
Untreated Observations: 662
***Notes*: Propensity scores for matching are generated by hazard regression (1) on Table I-2.**

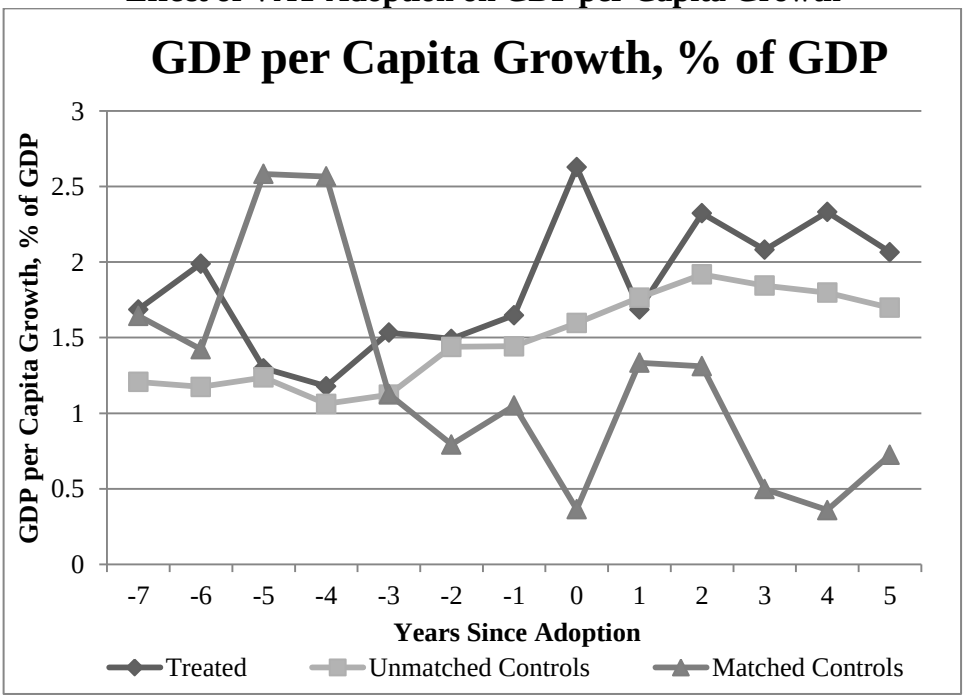
Appendix I-D: Additional Figures

Figure I-4
Effect of VAT Adoption on Investment



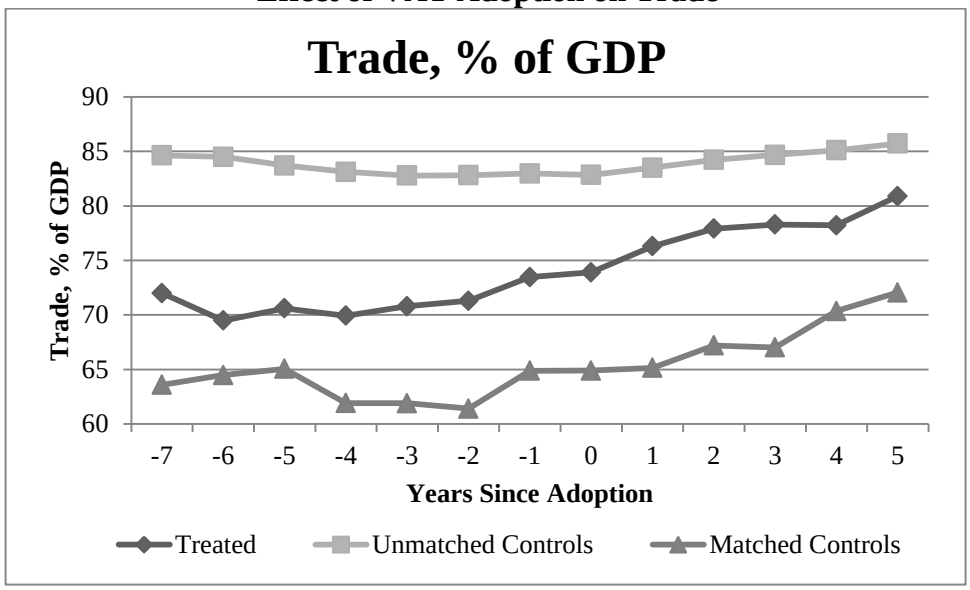
Notes: **VAT adopters are treated.**

Figure I-5
Effect of VAT Adoption on GDP per Capita Growth



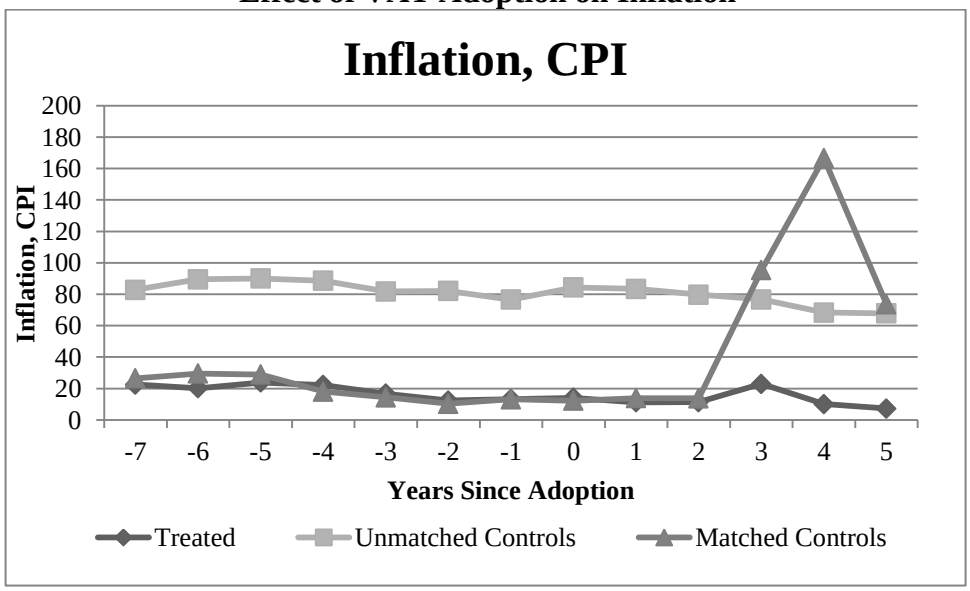
Notes: **VAT adopters are treated.**

Figure I-6
Effect of VAT Adoption on Trade



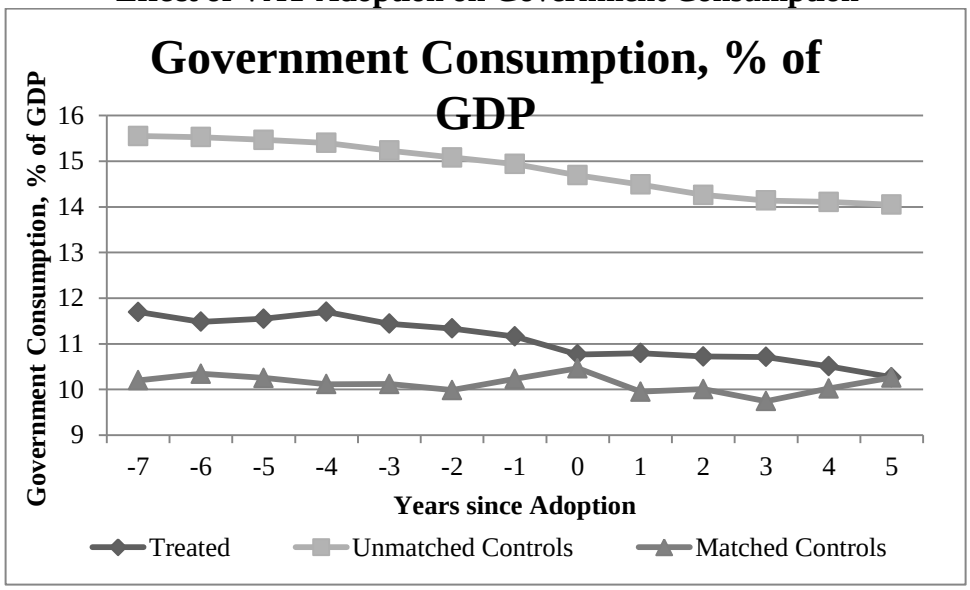
Notes: **VAT adopters are treated.**

Figure I-7
Effect of VAT Adoption on Inflation



Notes: **VAT adopters are treated.**

Figure I-8
Effect of VAT Adoption on Government Consumption



Notes: **VAT adopters are treated.**

Appendix I-E: Supporting Literature

Theoretical models show a VAT and other consumption taxes are less distortionary than other forms of taxation with respect to investment decisions, though nearly all taxes will distort such decisions to some degree (Van den Noord and Heady 2012, Zagler and Durnecker 2003.) Thus it is possible investment would be higher under a regime in which VATs are collected. Relevant to this prediction, a study on the investment behavior of VAT exempt firms in China finds no increases in investment for the VAT exempt firms compared to the non-exempt firms (Cai and Harrison 2012.) This could imply the presence of a VAT did not affect their investment decisions, as its removal had no effect. It could also point to other policy implementation problems. The effects of VATs on investment decision are thus not clear.

Due to a need for businesses to report purchases and sales in order to be taxed, a VAT may encourage tax evasion activities, though it is not clear whether increased effort to hide economic activity will actually result in more hidden activity, as VATs are in general harder to evade than a standard sales tax (e.g. Barbone et al 2012, Ebeke 2012, Keen 2010.) If activity is hidden in the informal sector, observed GDP could fall for two reasons. First activity that exists may be hidden, and thus not officially counted. Secondly, in an effort to hide activities, suboptimal decision are made, for example, investing in favor of easier to conceal or more mobile capital, further lowering GDP growth (Emran and Stiglitz 2005.) As a result, VATs could either raise economic growth due to less deadweight loss in collecting public funds and less evasion activities, or lower it due to increasing incentives for evasion activities. It is possible these

incentives, and thus direction of growth effect, are very different depending on the countries' level of development (Stiglitz 2008.)

As VATs are usually zero rated on exports, with producers paying no taxes on goods that are exported, and levied on imports at their full value, there are likely to be impacts on trade from passing a VAT. Nicholson (2012) finds that when countries pass a VAT they increase their trade with the United States, which may be indicative of increased trade in general from VAT passage. VATs are often replacements for simpler taxes on trade, and are often more amenable to exporters due to the reimbursement mechanism (Bird 2010.) However, because one is dependent on receiving timely rebates from the government for taxes paid, countries with poorly functioning civil services may not reimburse exporters and thus discourage, rather than encourage, exports (Hines and Desai 2005, Kelly and Fox 2006.)

VATs often replace turnover or production taxes, a tax on the value of a good at each stage of the manufacturing process. Since many of these taxes do not have mechanisms for rebates like the VAT does, this tax would be assessed multiple times throughout the process, with the end product's price being significantly inflated before reaching the consumer (Ebrill et al. 2001.) The VAT's rebate system, at least in theory, prevents this cascading of prices due to taxation. It is therefore possible that replacing some taxes with VATs will lower price pressures and reduce inflation that resulted from the structure of the previous taxes.

Finally, given the ease of tax collection, one may expect VAT adoption to lead to governments choosing to collect more taxes in order to raise spending. This fear of a "money machine" leads many, especially in the United States, to dislike the idea of a

VAT due to potential of government growth (Keen 2007.) However, it is hard to demonstrate if a VAT causes government to get larger without need for intent, or if a government adopts it in hopes of increasing spending. While Keen (2007) shows that VATs should, on its own, raise government spending if distortions fall, it is hard to see its net effect as governments may shift away from other sources of revenue that are more distorting.

Appendix I-F: Choice of Hazard model over Probit

This paper settled on modeling the VAT adoption decision via survival analysis in which countries exit estimation once they adopt the policy, a standard feature of survival analysis. If one were to use a probit to model whether or not a country has a VAT in a given period, it would require a dummy for having a VAT in previous periods, as because of the policies' persistence, this lagged VAT dummy will likely explain the majority of variation in the data for the countries that have adopted, as it did in Keen and Lockwood (2010.)

However, as the goal of this adoption equation is to create matches to evaluate the impact of VATs, one cannot have an overly deterministic adoption equation. If the adoption equation assigns observational units to treatment and control groups too strongly, accurate matching will be impossible as the distribution of propensity scores will be bimodal (Glick et al 2006.) The countries that do not adopt will cluster close at a very low propensity score, and the countries that do adopt will cluster at a very high propensity score. Accurately matching a treated (adopter) to a control (non-adopter) will be difficult as their propensity scores, and possibly underlying covariates, will be fundamentally different, and some overlap is necessary for matching to be successful. It will thus be unclear as to what is actually causing the difference in observed outcomes- the treatment itself, or the pre-existing differences between the groups choosing to be treated and not treated.

One of these two limitations cannot be resolved without exacerbating the other when modeling this phenomenon using a probit regression. To properly model the decision to have a VAT, one must include a lagged dummy for having a VAT in the

selection equation. To conduct matches, one must have some overlap in the distributions of propensity scores. If the lagged dummy for having a VAT causes assignment to treatment and control groups to become too deterministic, as is likely with only 5 repeal events, there will be no overlap of propensity scores and matching is thus impossible. If the lagged dummy is omitted, the probit will be mis-specified. Thus to have a probit that produces valid matches, one would have to omit a lagged policy dummy, but to properly model the decision, one must have a lagged policy dummy.

Another concern is that a country will be choosing to have a VAT multiple times under a probit specification, so it will be hard to separate the immediate effects of VAT adoption compared to longer term effects. Thus, I chose survival analysis to model this adoption decision. One could restrict the probit so one drops countries from estimation as soon as they adopt the policy, but this will produce results similar to a survival model. Coile and Gruber (2007) encounter a similar issue modeling the decision to retire for the first time and settle on using a probit restricted in such a way; results do not differ significantly.

Appendix I-G: Supporting Information on Matching

There are two main considerations that must be accounted for using matching in this paper in particular. First is the increasing scarcity of matches over time, a symptom of the overwhelming success of the VAT. The majority of countries studied adopts a VAT and as a result exits survival estimation at some point. In later years, being forced to match between adopters and only non-adopting countries in the same year could result in particularly poor propensity score matches as few control countries will remain. To overcome this problem, this paper puts no restrictions on country matching with regards to the time dimension, treating matching and control countries as one large cross section.[32] For example, the pairing that minimizes $|Xb_1-Xb_2|$, Jordan adopting in 2001 with a propensity score of 8.05 is Ethiopia not adopting in 1991 with a propensity score of 8.04.

Second is the possibility of contamination. As this paper seeks to look at the effect of VAT adoption over time, it will be necessary to look at the time path of variables of interest for several years after the adoption event. It is possible that a treatment country-year that adopts will be matched to a control country-year that adopts a VAT a few years after being matched. Trying to observe the long run impact of a VAT will thus be complicated, as one must choose a control country-year that does not adopt for several years. This problem should be common, as the two countries being matched will, by design, have very similar failure times for choosing a VAT. Failing to

[32] This would be a problem if one was systematically matching treatment country-years to controls country-years in a fundamentally different period. When comparing matches, this paper will look at mean differences in independent variables to see if there is a systematic difference in control and treatment observations, and no such systematic differences exist. Were there a systematic difference, there would be cause for concern.

control for this problem will downwardly bias results, as one will be comparing a treated country to a control that actually later received treatment. Referring to the earlier match, Spain 1986 should not be matched to Japan 1987, as Japan would adopt a VAT in 1989. If it were, in comparing the impacts of a VAT for 5 years after adoption, Spain 1991 to Japan 1992, the control would have actually received the treatment in that 5 year window and mean differences will be biased downward.

The most obvious fix carries with it a tradeoff. If one wishes to observe the effect of VAT for N years after it is first adopted, one can simply drop all countries as possible controls for the N years before that country adopts a VAT. As N gets very small, matches should not be adversely affected, but one may miss longer term effects of VAT adoption. As N grows larger, the matches will necessarily become worse as one is requiring the difference in survival times for treatment and control groups to be higher, but longer term effects can be studied. However, looking at very long term effects may be beyond the power of this model to begin with, as it matches on a single year and the key assumption that the only difference between treatment and control group is the treatment becomes less supportable on a longer time horizon. For the purpose of this paper, I will look at only a 5 year window after the VAT adoption event. This is the lower bound predicted by Bird et al (2005) on how long it will take to see the full effects of a VAT.

Appendix II-A: Countries Included

I) Countries Included in Table II-2 (92):

Albania; Algeria; Angola; Australia; Bahrain; Bangladesh; Benin; Bhutan; Botswana; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Central African Republic; Chad; China; Comoros; Congo, Dem. Rep.; Congo, Rep.; Croatia; Cyprus; Djibouti; Egypt, Arab Rep.; El Salvador; Eritrea; Ethiopia; Fiji; Finland; Gabon; Ghana; Guinea; Guinea-Bissau; Guyana; India; Iran, Islamic Rep.; Japan; Jordan; Kenya; Lao PDR; Lesotho; Liberia; Libya; Macedonia, FYR; Madagascar; Malawi; Malaysia; Mali; Mauritania; Mauritius; Mongolia; Morocco; Mozambique; Namibia; Nepal; New Zealand; Niger; Nigeria; Oman; Pakistan; Papua New Guinea; Paraguay; Philippines; Qatar; Romania; Rwanda; Saudi Arabia; Serbia; Sierra Leone; Singapore; Slovenia; Solomon Islands; South Africa; Sri Lanka; Sudan; Suriname; Swaziland; Switzerland; Syrian Arab Republic; Tanzania; Thailand; Togo; Trinidad and Tobago; Tunisia; Uganda; United Arab emirates; United States; Venezuela, RB; Vietnam; Yemen, Rep.; Zambia; Zimbabwe.

II) Countries included in Matching

Note the countries included in each matching specifications will be slightly different in each case, as the method here demands 13 continuous years of data (7 before Year of Adoption and 5 after plus the actual year of the adoption.) Thus, one can only study the impact of adoptions from 1986-2007, even though one can study the determination of adoptions up through 2010. Exact countries included in each specification are available upon request from the author.

Appendix II-B: Data Definitions

Agriculture: Percent of GDP contributed by agricultural activities, from the World Development Indicators, updated January 2015.

Bordering Countries with VAT: Number of bordering countries with a VAT divided by the number of bordering countries, from author's calculations updated February 2015.

British Commonwealth: Equal to 1 if a country is a member of the British Commonwealth at any time in the sample period, from the Commonwealth Network updated February 2015.

Exports: Exports as a percent of GDP, from the World Development Indicators updated January 2015.

Federal: Equal to 1 if a country is governed as a federation in that year, from Treisman (2002) with updates from the author, updated February 2015.

Francophonie: Equal to 1 if a country is a member of the Francophonie at any time in the sample period, from Organisation internationale de la Francophonie updated February 2015.

GDP Per Capita: Natural log of GDP per capita as measured by current US dollars, from the World Development Indicators updated January 2015.

Government Consumption: Government Consumption, as a percent of GDP from the Penn World Tables updated September 2013.

IMF Lending Dummy: Equal to 1 if the country has a nonzero balance in either a PRG or GRA credit facility in any months of that year, a stock measure, from the IMF Financial Data query tool updated February 2015.

Imports: Imports as a percent of GDP, from the World Development Indicators updated January 2015.

Inflation: The natural log of 100 plus the yearly inflation as measured by the CPI, from the World Development Indicators updated January 2015.

Institutions: The polity2 variable of level of democracy as measured by the Polity IV project updated February 2015.

Investment: Gross capital formation as a percent of GDP from the World Development Indicators updated January 2015.

Island: Equal to 1 if a country has no other countries it borders by land, from author's calculations updated February 2015.

Landlocked: Equal to 1 if a country has no land bordering an ocean, from author's calculations updated February 2015.

MF External Debt: Externally held debt as a share of GDP as measured by Lane and Milesi-Feretti updated 2011.

Natural Resource Rents: Percent of GDP coming from natural resource production, from the World Development Indicators updated January 2015.

Population: Natural log of population, from the World Development Indicators updated January 2015.

% of Population 65+: Percent of population above 65 years of age, from the World Development Indicators updated January 2015.

Size, Square Kilometers: Natural log of the surface area of a country, from the World Development Indicators updated January 2015.

Warsaw Pact: Equal to 1 if a country was a Warsaw Pact Nation, from author's calculations updated February 2015.

WDI External Debt: Externally held debt as a share of GDP as recorded by the World Development Indicators updated January 2015.\

WEO Central Government Debt: central government debt as a share of GDP as recorded by the World Economic Outlook updated February 2014.

WEO Expenditure: government expenditure as a share of GDP as recorded by the World Economic Outlook updated February 2014.

WEO Primary Surplus or Deficit: Cyclically adjusted surplus or deficit as a share of GDP as recorded by the World Economic Outlook updated February 2014.

WEO Surplus or Deficit: Surplus or deficit as a share of GDP as recorded by the World Economic Outlook updated February 2014.

WEO Tax Revenue: Tax revenue as a share of GDP as recorded by the World Economic Outlook updated February 2014.

Appendix III-A: Countries Included

Countries with at least one observation of TOEFL scores for the years studied (173):

Afghanistan; Albania; Algeria; American Samoa; Andorra; Angola; Argentina; Armenia; Aruba; Australia; Austria; Azerbaijan; Bahrain; Bangladesh; Belarus; Belgium; Benin; Bhutan; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Brunei Darussalam; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Republic; Chad; Chile; China; Colombia; Congo, Dem. Rep.; Congo, Rep.; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Republic; Denmark; Djibouti; Dominican Republic; Ecuador; Egypt, Arab Rep.; El Salvador; Eritrea; Estonia; Ethiopia; Fiji; Finland; France; French Polynesia; Gabon; Gambia, The; Georgia; Germany; Ghana; Greece; Guatemala; Guinea; Guinea-Bissau; Haiti; Honduras; Hong Kong SAR, China; Hungary; Iceland; India; Indonesia; Iran, Islamic Rep.; Iraq; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea, Dem. Rep.; Korea, Rep.; Kosovo; Kuwait; Kyrgyz Republic; Lao PDR; Latvia; Lebanon; Lesotho; Liberia; Libya; Lithuania; Luxembourg; Macao SAR, China; Macedonia, FYR; Madagascar; Malawi; Malaysia; Mali; Marshall Islands; Mauritania; Mauritius; Mexico; Micronesia, Fed. Sts.; Moldova; Monaco; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Caledonia; New Zealand; Nicaragua; Niger; Nigeria; Northern Mariana Islands; Norway; Oman; Pakistan; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Puerto Rico; Qatar; Romania; Russian Federation; Rwanda; San Marino; Saudi Arabia; Senegal; Sierra Leone; Singapore; Slovak Republic; Slovenia; Somalia; South Africa; Spain; Sri Lanka; Sudan; Suriname; Swaziland; Sweden; Switzerland; Syrian Arab Republic; Tajikistan; Tanzania; Thailand; Togo; Tonga; Tunisia; Turkey; Turkmenistan; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States; Uruguay; Uzbekistan; Venezuela, RB; Vietnam; Yemen, Rep.; Zambia; Zimbabwe.

Appendix III-B: Data Definitions

Africa: 1 if area is in Africa, as defined by ETS.

Americas: 1 if area is in the Americas, as defined by ETS.

Asia: 1 if area is in Asia, as defined by ETS.

Emigration, % of Population: OECD immigration data showing how large the expatriate population of a country is relative to the total population of that origin country, updated 2008.

Ethnolinguistic Fractionalization: Greener sum of squares index constructed by Roeder 2001, comparing whether two randomly selected individuals in a country share the same mother tongue.

Europe: 1 if area is in Europe, as defined by ETS. Omitted in most specifications.

FDI, % of GDP: Net foreign investment as a share of GDP from World Development indicators.

GDP Per Capita, Ln real 2005 US dollars: Natural log of GDP per capita as measured in Real 2000 US dollars, from the World Development Indicators.

Government Spending, % of GDP: Government consumption as a share of GDP from Penn world tables.

Hours to Learn English: How many hours for a native English speaker to learn the official languages of a given country. This is assumed to be symmetric, so it would take the same time for the native official language speaker to learn English. From Foreign Service Institute.

Inflation, GDP Deflator: Inflation as measured by the GDP deflator, from the World Bank.

Institutional Quality: Sum of World Government indicators from the World Bank.

Investment, % of GDP: Gross capital formation as a share of GDP, from the World Bank.

Island: Equal to 1 if a country has no other countries it borders by land, from author's calculations.

Land Area, Ln: Natural log of square kilometers of land area of a country.

Landlocked: Equal to 1 if a country has no land bordering an ocean, from author's calculations.

Latitude: The latitude of a country from CIA World Factbook.

Linguistic Proximity-2: Measure of linguistic proximity with higher values being closer to English, from Melitz and Toubal (2014.)

Middle East: 1 if area is in the Middle East, as defined by ETS.

Official English: Equal to 1 if a countries' official languages include English, CIA world Factbook.

Population: Natural Log of population, from the World Bank

Pacific: 1 if area is in the Pacific, as defined by ETS.

Schooling 1990: Average years of education per person in 1990, from Barro-Lee.

Speaks English, % of Population: % of the population that reports they are Proficient in English. Provided from various national census bureaus and Crystal (2003)

TOEFL Score: Test of English as a Foreign Language, the Average TOEFL score in a country, from ETS 1992-2012 omitting 2011 as data is unavailable. TOEFL scores are rescaled here so that paper, computer and internet based test scores can be compared. Paper scores are rescaled such that the minimum score in the first year available for the paper exam, 223 in 1992, is the same as the lowest score in the first year recorded for the internet based exam, 0 in 2006, and so that the maximum scores are the same, and 677 and 120 respectively. This yields a (rounded) linear transformation of $Y = (X-223)/3.78$, where X is the paper score and Y is the comparable transformed score. Using similar methodology with a reference year of 1998, $Y = (X–7)/2.44$ is used to transform computer test scores into internet based test scores using a base year of 1998 with a minimum score of 7 and maximum score of 300.